

SILVIO JACKS DOS ANJOS GARNÉS

**RESOLUÇÃO DAS AMBIGÜIDADES GPS PARA LINHAS DE BASE
CURTA: ANÁLISE DOS ALGORITMOS DE OTIMIZAÇÃO**

**Tese apresentada como requisito parcial à
obtenção do título de Doutor em Ciências,
Curso de Pós-Graduação em Ciências
Geodésicas, Setor de Ciências da Terra,
Universidade Federal do Paraná.**

Orientação:

Prof. Dr. Raimundo José Borges de Sampaio

Prof. Dr. Quitino Dalmolin

Prof.^a Dr.^a Claudia Pereira Krueger

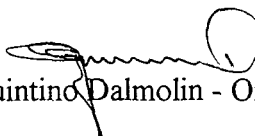
**CURITIBA
2001**

**“RESOLUÇÃO DAS AMBIGÜIDADES GPS PARA LINHAS DE
BASE CURTA: ANÁLISE DOS ALGORITMOS DE
OTIMIZAÇÃO”**

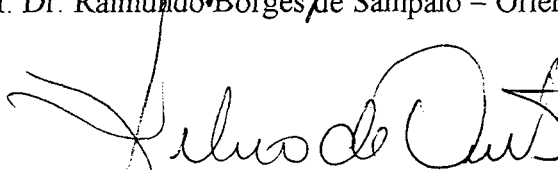
POR

SÍLVIO JACKS DOS ANJOS GARNÊS

Tese nº 022 aprovada como requisito parcial ao grau de Doutor no Curso de Pós-Graduação em Ciências Geodésicas da Universidade Federal do Paraná, pela Comissão formada pelos professores:



Prof. Dr. Quintino Dalmolin - Orientador e Presidente (UFPR)


Prof. Dr. Raimundo Borges de Sampaio – Orientador (PUC)


Prof. Dr. Sílvio Rogério Correia de Freitas - Membro (UFPR)


Prof. Dr. Denizar Blitzkow - Membro (EPUSP/PTR)


Prof. Dr. João Francisco Galera Monico – Membro (UNESP)


Prof. Dr. Luiz Danilo Damasceno Ferreira – Membro (UFPR)

AGRADECIMENTOS

Externo meus agradecimentos em primeiro lugar a minha família: Maria Aparecida Vasconcelos dos Santos (Esposa); Rafael Vasconcelos Garnés (Filho); Matheus Vasconcelos Garnés (Filho); Silvia Vasconcelos Garnés (Filha), pelo incentivo diário e pela força para continuar vivendo.

À Universidade para o Desenvolvimento do Estado e da Região do Pantanal – UNIDERP, pela concessão da licença para minha qualificação e apoio através do seu plano de capacitação de docente.

Ao Professor Dr. Raimundo José Borges de Sampaio por ter sido muito mais do que um orientador, a quem devo grande parte dos meus conhecimentos e de minha personalidade acadêmica.

Ao Professor Dr. Quintino Dalmolin pelo grande ser humano que é e pelas várias correções e idéias sugeridas.

À Professora Dra. Claudia Pereira Krueger pelo companheirismo e ajuda, principalmente relacionada com a parte do GPS.

Ao Professor Dr. Silvio Rogério Correia de Freitas pelo seu apoio pessoal e também como Coordenador do Curso de Pós-Graduação em Ciências Geodésicas, nas dificuldades particulares passadas durante o ano de 2000.

À minha mãe, Etelvina dos Anjos, pelos conselhos e por arcar com algumas de minhas responsabilidades, favorecendo o meu desempenho psicológico.

Aos meus familiares que não mediram esforços em me ajudar quando precisei: Flávio, Rosangela, Cândida, Pedro, Sérgio, Daúva, Hideraldo, Mário e Sueli .

Aos vizinhos e amigos de Curitiba que se tornaram minha segunda família durante o tempo em que morei no Bairro Capão da Imbuia: Luiz, Sirley, Jean, Maristela, Paulo, Wilson, Astéria, Laura, Darci e Luiz Eduardo.

Aos demais amigos e colegas que de uma forma ou de outra me ajudaram e me apoiaram ao longo dessa pesquisa, peço-lhes desculpas por não citar seus nomes, pois não foram poucos e não gostaria de me sentir culpado esquecendo de alguém.

À Fundação Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – CAPES, pela concessão da bolsa de estudos.

A **Deus**, pela **vida** e pelo **dom**, sem os quais nada poderia ter sido realizado.

A Verdade

É a filosofia para todos os momentos da vida, não é “absoluta”, e não deve se desalentar aquele que descobre que sua crença de um momento para outro passou a ser uma verdade “relativa”. Assim, de verdade, em verdade, avança a humanidade no desvendar das leis naturais do Universo, tudo há seu tempo, conforme desígnios de uma Verdade Maior (Deus).

O Autor

SUMÁRIO

LISTA DE TABELAS E QUADROS	x
LISTA DE FIGURAS	xi
LISTA DE SÍMBOLOS	xii
RESUMO	xiii
ABSTRACT	xiv
1 INTRODUÇÃO	1
1.1 CONSIDERAÇÕES INICIAIS	1
1.2 DEFINIÇÕES DOS PROBLEMAS	4
1.3 OBJETIVOS	7
1.4 METODOLOGIA	8
1.5 ESTRUTURA DO TRABALHO	9
2 FUNDAMENTOS DO GPS	10
2.1 SEGMENTO ESPACIAL	10
2.1.1 Configuração Orbital e Satélites GPS	10
2.1.2 Sinais Transmitidos pelos Satélites	13
2.1.3 Degradação da Acurácia por Meio de Técnicas de Segurança	14
2.2 SEGMENTO DE CONTROLE	16
2.3 SEGMENTO DO USUÁRIO	17
2.3.1 Tipos de Receptores	17
2.3.2 Técnicas <i>Codeless</i> e <i>Quasi-codeless</i>	19
2.3.2.1 Quadratura do sinal (<i>signal squaring</i>)	19
2.3.2.2 Correlação cruzada (<i>cross-correlation</i>)	20
2.3.2.3 Quadratura após a correlação do código (<i>code-correlating squaring</i>)	20
2.3.2.4 <i>Z Tracking</i>	20
2.4 SISTEMA DE TEMPO GPS	21

2.5 PRINCIPAIS ERROS DO GPS.....	22
2.5.1 Erro na Posição Orbital	23
2.5.2 Erro do Relógio do Satélite.....	26
2.5.3 Relatividade	27
2.5.4 Atraso do <i>Hardware</i>	28
2.5.5 Erro entre Canais (<i>Interchannel Biases</i>).....	28
2.5.6 Erro do Relógio do Receptor	29
2.5.7 Ruído do Receptor.....	29
2.5.8 Centro de Fase.....	29
2.5.9 Refração Ionosférica.....	33
2.5.10 Refração Troposférica	38
2.5.11 Multicaminho	42
2.6 EQUAÇÕES DE OBSERVAÇÃO	45
2.6.1 Equação de Observação do Código.....	45
2.6.2 Equação de Observação da Fase da Portadora.....	48
2.6.3 Combinação Linear das Observáveis	52
2.6.3.1 Combinação linear das observáveis em L_1 e L_2	53
2.6.3.2 Equação de observação da simples diferença	55
2.6.3.3 Equação de observação da dupla diferença	57
2.6.3.4 Equação de observação da tripla diferença.....	58
3 AJUSTAMENTO DE OBSERVAÇÕES.....	60
3.1 PROBLEMA DE MÍNIMOS QUADRADOS	61
3.1.1 Problema de Mínimos Quadrados Linear	61
3.1.2 Mal-Condicionamento dos Sistemas de Equações Lineares	66
3.1.3 Métodos e Algoritmos para a Solução do Problema de Mínimos Quadrados Linear	71
3.1.3.1 Decomposição de Cholesky	72
3.1.3.2 Decomposição QR.....	74
3.1.3.3 Decomposição de valor singular - SVD	76

3.1.3.4 Método dos gradientes conjugados	78
3.1.4 Ajustamento Sequencial	85
3.1.5 Problema de Mínimos Quadrados Inteiro.....	87
3.1.5.1 O problema de mínimos quadrados inteiro.....	89
3.1.6 Problema de Mínimos Quadrados Linear com Injunções de Igualdade.....	92
3.1.7 Problema de Mínimos Quadrados Não-Linear	96
3.1.7.1 Método de Newton	96
3.1.7.2 Método de Gauss-Newton	98
3.1.7.3 Método de Levenberg-Marquardt	100
3.1.7.4 Método de Newton globalizado	101
3.1.7.5 Método de Gauss-Newton modificado globalizado	113
3.2 FUNDAMENTOS ESTATÍSTICOS	114
3.2.1 Esperança Matemática.....	115
3.2.2 Momentos, Variâncias e Covariâncias	116
3.2.3 Propagação das Covariâncias.....	119
3.2.4 Matriz dos Pesos para as Observáveis GPS.....	120
3.2.5 Caracterização Estatística para o Problema de Mínimos Quadrados.....	127
3.2.6 Testes de Hipóteses	131
4 RESOLUÇÃO DAS AMBIGÜIDADES GPS	135
4.1 MÉTODOS DE INICIALIZAÇÃO	135
4.1.1 Troca ou Intercâmbio de Antenas	136
4.1.2 Linha de Base Conhecida	137
4.2 MÉTODOS DE BUSCA NO ESPAÇO DAS COORDENADAS	142
4.2.1 Método da Função de Ambigüidade	142
4.3 MÉTODOS DE BUSCA NO ESPAÇO DAS AMBIGÜIDADES.....	148
4.3.1 Solução Baseada no Desvio Padrão das Ambigüidades.....	148
4.3.2 Fast Ambiguity Resolution Approach (FARA)	150
4.3.3 Busca com a Decomposição de Cholesky	153
4.3.4 Least-Squares Ambigüity Decorrelation Adjustment (LAMBDA)	155

4.3.5 Solução com Satélites Primários e Secundários	161
4.4 MÉTODOS INTEGRADOS	165
4.4.1 O Problema de Mínimos Quadrados Linear Misto	166
4.4.1.1 Métodos de funções de penalidade.....	167
4.4.1.2 Métodos de funções de penalidade sobre os não-inteiros.....	172
4.4.1.3 Solução do problema de mínimos quadrados linear misto por métodos de penalidade	174
4.4.2 Busca Direcional nos Espaços das Ambigüidades e das Coordenadas	176
5 CONCLUSÕES E RECOMENDAÇÕES	187
5.1 CONCLUSÕES	187
5.2 RECOMENDAÇÕES	191
REFERÊNCIAS BIBLIOGRÁFICAS	192
ANEXOS	198
ANEXO 1 - LINHA DE BASE (5,002 m): EXTRATO DE RESULTADOS.....	199
ANEXO 2 - LINHA DE BASE (185,153 m) : EXTRATO DE RESULTADOS	201
ANEXO 3 - LINHA DE BASE (950,565 m): EXTRATO DE RESULTADOS	203

LISTA DE TABELAS E QUADROS

TABELA 1.1	- NÍVEL DE RUÍDO DAS MEDIDAS	1
TABELA 2.1	- SITUAÇÃO DA CONSTELAÇÃO DE SATÉLITES GPS (ATUALIZAÇÃO/2001).....	12
QUADRO 2.1	- COMPARAÇÃO DAS TÉCNICAS USADAS PARA O AS.....	21
TABELA 2.2	- ÓRBITAS DO GPS	24
TABELA 2.3	- INFLUÊNCIA DO ERRO ORBITAL NO POSICIONAMENTO RELATIVO	25
TABELA 2.4	- ELEMENTOS DE CALIBRAÇÃO DA ANTENA.....	31
TABELA 2.5	- ELEMENTOS DE CALIBRAÇÃO DO CENTRO DE FASE.....	31
TABELA 2.6	- COMBINAÇÕES LINEARES DA FASE DAS PORTADORAS.....	54
TABELA 3.1	- SUBESPAÇOS FUNDAMENTAIS DA MATRIZ A	62
TABELA 4.1	- COORDENADAS DOS SATÉLITES E DAS ESTAÇÕES PARA UMA PRÉ-ANÁLISE DA TÉCNICA DA LINHA DE BASE CONHECIDA.....	140

LISTA DE FIGURAS

FIGURA 1.1 - INTERPRETAÇÃO GEOMETRICA DA AMBIGÜIDADE GPS	2
FIGURA 2.1 - POSICIONAMENTO ABSOLUTO SEM O δ - SA.....	16
FIGURA 2.2 - SISTEMA DE TEMPO GPS.....	22
FIGURA 2.3 - ERROS DO GPS	23
FIGURA 2.4 - ERRO NA POSIÇÃO ORBITAL X ERRO NA LINHA DE BASE	26
FIGURA 2.5 - DIAGRAMA DE REFERÊNCIA DA ANTENA TRIMBLE CHOKE RING.....	32
FIGURA 2.6 - CARACTERIZAÇÃO DO CENTRO DE FASE NA ESTAÇÃO PARA	32
FIGURA 2.7 - DISPERSÃO NA IONOSFERA	33
FIGURA 2.8 - ATMOSFERA SIMPLIFICADA	40
FIGURA 2.9 - REFRAÇÃO TROPOSFÉRICA	41
FIGURA 2.10 - MULTICAMINHO	42
FIGURA 2.11 - MEDIDA COM O CÓDIGO	46
FIGURA 2.12 - GEOMETRIA DA SIMPLES DIFERENÇA ENTRE RECEPTORES	55
FIGURA 2.13 - GEOMETRIA DA DUPLA DIFERENÇA.....	57
FIGURA 3.1 - INTERPRETAÇÃO GEOMÉTRICA DO PROBLEMA DE MÍNIMOS QUADRADOS LINEAR.....	62
FIGURA 3.2 - INTERPRETAÇÃO GEOMÉTRICA DO PROBLEMA DE MÍNIMOS QUADRADOS LINEAR PONDERADO.....	65
FIGURA 3.3 - NÚMERO DE CONDIÇÃO DA MATRIZ DE EQUAÇÕES NORMAIS DA DUPLA DIFERENÇA DE FASE.....	68
FIGURA 3.4 - INTERPRETAÇÃO GEOMÉTRICA DO PROBLEMA DE MÍNIMOS QUADRADOS INTEIRO.....	91
FIGURA 3.5 - CONDIÇÕES PARA CONVERGÊNCIA GLOBAL.....	102
FIGURA 4.1 - ESPAÇO DE BUSCA DAS COORDENADAS	146
FIGURA 4.2 - GRÁFICO DA FUNÇÃO DE AMBIGÜIDADE.....	147
FIGURA 4.3 - GRÁFICO DA FUNÇÃO DE PENALIDADE HIPERBÓLICA.....	171
FIGURA 4.4 - BUSCA INTEIRA PELA MÁXIMA DIREÇÃO DE DESCIDA	178
FIGURA 4.5 - ESPAÇO DE BUSCA DAS AMBIGÜIDADES: CASO BI-DIMENSIONAL	179

LISTA DE SÍMBOLOS

\supset	- Contém
R^n	- Espaço vetorial n-dimensional dos reais
Z^n	- Espaço vetorial n-dimensional dos inteiros
\perp	- Perpendicular
$\langle \mathbf{x}, \mathbf{y} \rangle$	- Produto interno dos vetores \mathbf{x} e \mathbf{y}
$\ \cdot \ $	- Norma de vetor ou norma de matriz, depende do contexto
\rightarrow	- Leva em
\Rightarrow	- Implica
$\Re(\mathbf{A})$	- Espaço coluna da matriz \mathbf{A}
$\aleph(\mathbf{A})$	- Espaço nulo da matriz \mathbf{A}
$C(\mathbf{A})$	- Número de condição da matriz \mathbf{A}
$\text{tr}(\mathbf{A})$	- Traço da matriz \mathbf{A}
\mathbf{A}^τ	- Pseudo-inversa da matriz \mathbf{A}
\mathbf{d}_k	- Direção do k-ésimo passo de um método iterativo
$\nabla f(\mathbf{x})$	- Vetor gradiente da função $f(\mathbf{x})$
$\nabla^2 f(\mathbf{x})$	- Matriz Hessiana da função $f(\mathbf{x})$
φ	- Fase da portadora, ou latitude, depende do contexto
λ	- Comprimento de onda, ou longitude, ou autovalor, depende do contexto
h	- Altitude geométrica
σ	- Desvio padrão de uma variável aleatória
σ^2	- Variância de uma variável aleatória
u	- Unidade de arredondamento do computador
Σ	- Matriz variância-covariância, ou matriz de valores singulares da decomposição de valor singular, depende do contexto
$E\{ \}$	- Esperança matemática de uma variável aleatória
$\sigma_0^2, \hat{\sigma}_0^2$	- Variância de referência a priori e a posteriori, respectivamente
$\tilde{\sigma}_0^2$	- Variância de referência a posteriori com as ambigüidades injuncionadas
v	- Graus de liberdade
α	- Nível de significância de um teste estatístico
H_0, H_1	- Hipótese nula e hipótese alternativa de um teste estatístico, respectivamente
$\prod_{i=1}^t n_i$	- Produto dos elementos n_i , $i=1, \dots, t$

RESUMO

Nesse trabalho é feita uma completa análise dos métodos de minimização do problema de mínimos quadrados linear e não-linear, no ajustamento de observações GPS, e de métodos de resolução das ambigüidades GPS para linhas de base curta. O trabalho aborda os mais sofisticados métodos, diretos e iterativos, atualmente disponíveis, para a solução do problema de mínimos quadrados. Esses métodos são aplicados e analisados quando a matriz de coeficientes dos parâmetros tem número de condição elevado, e o mal-condicionamento do sistema é uma realidade. Também, foram desenvolvidos e analisados métodos de minimização com convergência global. O primeiro, baseando-se em região de confiança, com compromisso de Levenberg-Marquardt e, o segundo, com direção de descida definida pela decomposição de Cholesky modificada, cuja minimização direcional utiliza-se *backtracking*. Para a resolução das ambigüidades GPS, desenvolveu-se um método por meio da solução do problema de mínimos quadrados linear misto, usando-se métodos de penalidade com as funções de barreira e de penalidade tradicionais e, também, com a função de penalidade hiperbólica. Além disso, fez-se uso de uma função de penalidade sobre os não-inteiros para obrigar a convergência sobre os inteiros. Desenvolveu-se além desses, um método integrado baseado-se em busca direcional relacionando o espaço das ambigüidades e o espaço das coordenadas. A direção de busca em tal método, preferencialmente, é definida pela direção do semi-eixo maior do hiper-elipsóide que define o espaço das ambigüidades, cuja relação, é estrita, com a direção do semi-eixo maior do elipsóide que define o espaço das coordenadas. A busca da solução envolve, ainda, grupos de satélites primários e secundários junto com a aplicação da função de ambigüidade. Experiências realizadas em linhas de base curta, e medidas da fase da portadora em L_1 são usadas para verificação dos principais aspectos dos métodos apresentados.

Palavras-chave: Ambigüidades; GPS; Ajustamento de observações.

ABSTRACT

In this work we carry out a full analysis in the methods of minimization for linear and non-linear least squares problems, in the adjustment of the GPS observations, and for GPS ambiguity resolution methods in short base line. This work pointed out the more sophisticated direct and iterative methods currently available for the solution of the least squares problems. These methods are applied and analyzed when the design matrix of the parameters has a very high condition number, and the linear system is ill-conditioning. It is also developed and analyzed methods for minimization of global convergence, as well. The first approach is based on trust region, using Levenberg-Marquardt compromise and, the second approach using Choleky modified decomposition for the decent direction, with linear search and backtracking. For the GPS ambiguity resolution it is carried out a method which use the solution of the mixed linear least square problem using application penalty methods with penalty and barrier functions. It was also used hyperbolic penalty function. The development of these applied techniques were integrated upon method based in the directional search between ambiguity space and coordinates space. The search direction is preferentially that one defined by the major semi-axis of the ellipsoid hipper-dimensional that define the ambiguity space, which is related to the strict direction of the major semi-axis of the ellipsoid that define the coordinates space. The search of satellites include groups of primary satellites and secondary satellites, and the ambiguity function. Experiences were carried out in short base line whit L_1 carrier phase observations showing the aspects principals of the methods.

Key-words: Ambiguity; GPS; Adjustment of observations.

1 INTRODUÇÃO

1.1 CONSIDERAÇÕES INICIAIS

Após o lançamento do satélite artificial Russo, SPUTNIK I, em 4 de outubro de 1957, surgiram as primeiras idéias de elaborar um sistema de navegação baseado em ondas de rádio com transmissão via satélite. Explorando o efeito Doppler, a Marinha dos Estados Unidos implementou o programa TRANSIT (com três satélites protótipos) que passou a operar a partir de 1964. A partir 1967 o programa foi expandido e passou a se chamar *Navy Navigation Satellite System* – NNSS. Na busca de um sistema de navegação que fosse ainda mais rápido e efetivo, iniciou-se em 1973, cotejo dos projetos TIMATION da Marinha Americana e SYSTEM 621-B da Força Aérea, o desenvolvimento do projeto *Navigation System with Time and Ranging Global Positioning System* – NAVSTAR GPS (ANDRADE, 1988, p. 1-2).

Inicialmente o GPS (sigla pelo qual o NAVSTAR GPS ficou mundialmente conhecido) deveria trabalhar apenas com medidas de distâncias receptor(antena)/satélites(antenas) derivadas a partir do tempo para a máxima correlação dos códigos *Clear/Acquisiton* (C/A) e/ou *Precision* (P). Todavia, as distâncias receptor/satélites com os códigos, só são obtidas com acurácia na ordem do decímetro (ver ruído na tab.1.1) e para aplicações geodésicas, exige-se observações mais refinadas. Tais observações são proporcionadas por medidas da fase da portadora (tab. 1.1).

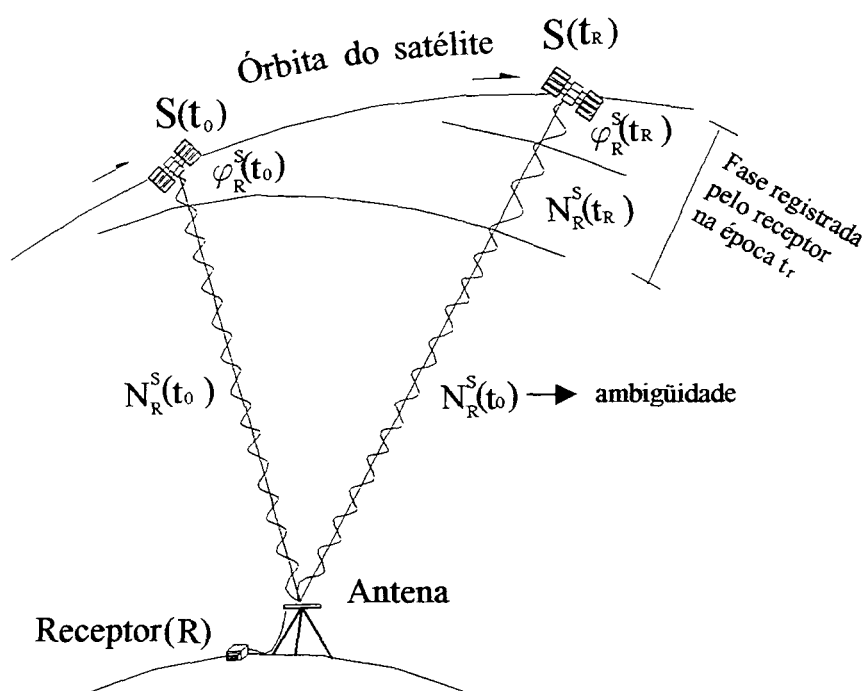
TABELA 1.1 – NÍVEL DE RUÍDO DAS MEDIDAS

OBSERVÁVEL	FREQÜÊNCIA MHz	COMPRIMENTO DE ONDA	RUÍDO		
C/A	1,023	293 m	10 cm	—	3 m
P	10,23	29,3 m	10 cm	—	30 cm
L ₁	1.575,42	19,05 cm	0,5 mm	—	5 mm
L ₂	1.227,60	24,45 cm	0,5 mm	—	5 mm

FONTE: Adaptada de HOFMANN-WELLENHOF, B.; LICHTENEGGER, H.; COLLINS, J. **Global positioning system: Theory and practice**. 3. ed. New York: Springer-Verlag, 1994, p.94.

O inconveniente em se utilizar as observações da fase em comparação com as do código, é que o receptor registra somente os ciclos inteiros acumulados a partir da primeira observação, mais a parte fracionária da fase medida pelo comparador. Assim, fica incógnito o número inteiro de ciclos desde o centro de fase da antena do satélite, até o centro de fase da antena do receptor (no instante da primeira observação). A este número inteiro de ciclos denominou-se *ambigüidade GPS*. O conceito geométrico está ilustrado na fig.1.1.

FIGURA 1.1 - INTERPRETAÇÃO GEOMÉTRICA DA AMBIGÜIDADE GPS



FONTE: Adaptada de HOFMANN-WELLENHOF, B.; LICHTENEGGER, H.; COLLINS, J. **Global positioning system: theory and practice**. 3 ed. New York: Springer-Verlag, 1994, P. 92.

NOTA: Para transformar unidades de ciclos em unidades lineares, basta a multiplicação pelo respectivo comprimento de onda da portadora.

A importância de se ter o valor correto das ambigüidades fica evidenciado pela análise da fig. 1.1 e tab. 1.1, pois, é como se a distância entre receptor/satélite pudesse

ser medida com precisão milimétrica, a menos é claro, de alguns erros como: refração atmosférica, estados dos relógios do receptor e satélite em relação ao tempo GPS e outros (ver seção 2.5).

É difícil modelar todos os erros envolvidos nas observações da distância entre receptor/satélite e, qualquer resíduo sistemático resultante afeta a determinação correta das ambigüidades. Entretanto, quando o posicionamento é relativo, muitos erros em comum são cancelados (ou minimizados a quantidades negligenciáveis), a natureza inteira das ambigüidades (combinação linear das originais) é mantida e elas podem ser separadas dos demais parâmetros envolvidos.

Uma vez resolvidas às ambigüidades¹, a cada nova época observada (desde que não ocorra perda de ciclos), o posicionamento passa a ter precisão centimétrica.

Quando um receptor GPS possui a capacidade de rastrear as fases das portadoras L_1 e L_2 , existe a possibilidade de realizar alguns tipos de combinações lineares entre elas, de maneira que a observável resultante, tenha um comprimento de onda maior do que as originais (ver seção 2.6.3.1) e, com isto, o problema de resolver as ambigüidades passa a ser bem mais simples.

Embora os métodos desenvolvidos nesse trabalho possam ser aplicados com sucesso em observáveis formadas da combinação linear das duas portadoras, o principal enfoque é considerar apenas uma portadora. Os motivos são:

- i) receptores de dupla frequência são muito mais caros do que os de uma frequência, e não estão acessíveis para a realidade econômica da maioria dos usuários que utilizam esta tecnologia, pelo menos no Brasil;
- ii) grande parte das aplicações práticas são relativas e, também, podem ser realizadas em linhas de base curta, ou senão, podem ser restringidas a elas. Nestas linhas, os efeitos sistemáticos (refração

¹ As ambigüidades são ditas resolvidas para uma determinada observável, quando o número inteiro de ciclos é determinado corretamente.

atmosférica da dupla diferença) principalmente devido a ionosfera são bastante reduzidos e, a não ser que a aplicação justifique, a modelagem com ambas as frequências não é necessária. Além disso, pode-se lançar mão de modelos regionais da ionosfera para corrigir tal efeito (ver modelos regionais da ionosfera em CAMARGO, 1999).

Na finalidade de resolver as ambigüidades GPS, em geral, é preciso realizar um ajustamento de observações da fase da portadora, por causa das observações em superabundância serem incompatíveis com o sistema de equações formado com um determinado modelo funcional.

Apesar de nos levantamentos cinemáticos e em tempo real o ajustamento recursivo por filtro de Kalman ser bastante utilizado, o critério objeto de investigação nesse trabalho, será apenas o de mínimos quadrados.

Os problemas a serem tratados envolvendo o ajustamento das observações e a resolução das ambigüidades, são definidos na seção (1.2) a seguir.

1.2 DEFINIÇÕES DOS PROBLEMAS

A observação GPS pode ser relacionada a parâmetros incógnitos através de uma função não-linear explícita, $f: R^n \rightarrow R$, conhecida por equação de observação:

$$l_b = f(\mathbf{x}) , \quad (1.1)$$

onde:

$l_b \in R$: representa a observação GPS;

$\mathbf{x} \in R^n$: representa o vetor dos parâmetros a estimar, cujas componentes são parte reais e parte inteiras, se as ambigüidades forem consideradas (lembrando que $R \supset Z$) .

Rastreando-se os sinais de mais de 4 satélites, dependendo do modelo (f), em 2 épocas já se tem um sistema superabundante de equações de observação:

$$\mathbf{L}_b = \mathbf{F}(\mathbf{x}), \quad (1.2)$$

onde:

\mathbf{L}_b : vetor m-dimensional de observações GPS;

$\mathbf{F} : \mathbf{R}^n \rightarrow \mathbf{R}^m$ não-linear em \mathbf{x} , $m > n$;

\mathbf{x} : vetor n-dimensional dos parâmetros.

Por causa da propriedade aleatória das observações (ver seção 3.2) e de algum fenômeno físico não parametrizado ou de algum erro grosseiro, a igualdade em (1.2) não se cumpre em geral. A diferença pode ser representada por uma função residual:

$$\mathbf{r}(\mathbf{x}) = \mathbf{F}(\mathbf{x}) - \mathbf{L}_b, \quad (1.3)$$

sendo:

$\mathbf{r} : \mathbf{R}^n \rightarrow \mathbf{R}^m$ não-linear em \mathbf{x} .

i) Problema de mínimos quadrados não-linear

É definido para minimizar a função residual:

$$\underset{\mathbf{x} \in \mathbf{R}^n}{\text{minimizar}} \quad \frac{1}{2} \|\mathbf{r}(\mathbf{x})\|_{\mathbf{Q}_{L_b}^{-1}}^2, \quad (1.4)$$

onde:

$\|(\cdot)\|_{\mathbf{Q}_{L_b}^{-1}} = \sqrt{(\cdot)^T \mathbf{Q}_{L_b}^{-1} (\cdot)}$: define uma norma de vetor;

$\mathbf{Q}_{L_b}^{-1}$: é uma matriz simétrica e definida positiva de pesos.

ii) Problema de mínimos quadrados linear

É definido para minimizar a função quadrática residual a partir da aproximação afim da função residual, como:

$$\underset{\mathbf{x} \in \mathbb{R}^n}{\text{minimizar}} \quad \frac{1}{2} \left\| \mathbf{v}(\mathbf{x}) \right\|_{\mathbf{Q}_{\mathbf{L}_b}^{-1}}^2, \quad (1.5)$$

onde:

$$\mathbf{v}(\mathbf{x}) = \mathbf{r}(\mathbf{x}_0) + \left. \frac{\partial \mathbf{r}}{\partial \mathbf{x}} \right|_{\mathbf{x}_0} (\mathbf{x} - \mathbf{x}_0),$$

sendo:

$$\mathbf{r}(\mathbf{x}_0) = \mathbf{F}(\mathbf{x}_0) - \mathbf{L}_b = -\mathbf{L} : \text{resíduo calculado no ponto } \mathbf{x}_0;$$

$$\left. \frac{\partial \mathbf{r}}{\partial \mathbf{x}} \right|_{\mathbf{x}_0} = \left. \frac{\partial \mathbf{F}}{\partial \mathbf{x}} \right|_{\mathbf{x}_0} = \mathbf{A} : \text{matriz Jacobiano calculada no ponto } \mathbf{x}_0.$$

iii) Problema de mínimos quadrados inteiro

É definido para minimizar a função quadrática:

$$\underset{\mathbf{x}_2 \in \mathbb{Z}^t}{\text{minimizar}} \quad (\mathbf{x}_2 - \hat{\mathbf{x}}_2)^T \mathbf{Q}_{\hat{\mathbf{x}}_2}^{-1} (\mathbf{x}_2 - \hat{\mathbf{x}}_2), \quad (1.6)$$

onde:

t : número de parâmetros a estimar;

\mathbf{x}_2 : vetor de parâmetros inteiros a estimar (ambigüidades);

$\hat{\mathbf{x}}_2$: vetor de parâmetros estimados sobre o \mathbb{R}^t ;

$\mathbf{Q}_{\hat{\mathbf{x}}_2}$: matriz cofatora do vetor estimado $\hat{\mathbf{x}}_2$.

iv) Problema da função de ambigüidade

Aborda a resolução das ambigüidades sob um outro aspecto. Nesse problema, a natureza inteira das ambigüidades não influencia no valor da função de ambigüidade (equação 1.7), que ao ser maximizada sobre o espaço das coordenadas, conduz à resolução das mesmas. O problema, numa forma simplificada (ver detalhes na seção 4.2.1), pode ser definido como:

$$\underset{\mathbf{x}_1 \in \mathbb{R}^3}{\text{maximizar}} \quad \left\| \sum_j e^{i\theta(\mathbf{x}_1)} \right\|, \quad (1.7)$$

onde:

\mathbf{x}_1 : vetor com as coordenadas da estação a determinar;

$i = \sqrt{-1}$: unidade imaginária;

$\theta(\mathbf{x}_1)$: ângulo no plano complexo.

Os problemas (i) e (ii) estão relacionados com as estimativas sobre os reais dos parâmetros, também chamadas estimativas da solução flutuante. Os problemas (iii) e (iv) estão relacionados com a resolução das ambigüidades em si, sendo que, o primeiro busca uma solução inteira para as ambigüidades sobre o espaço das ambigüidades e, o segundo busca uma solução real para as coordenadas sobre o espaço das coordenadas.

Ainda há um outro problema; o problema de mínimos quadrados linear misto, que será definido e tratado na seção (4.4.1). Nesse problema, a minimização da função objetivo é simultaneamente feita sobre o espaço das coordenadas e sobre o espaço das ambigüidades.

1.3 OBJETIVOS

Objetivo geral

Analisar e propor métodos e algoritmos para o ajustamento das observações

GPS, assim como, para a resolução das ambigüidades.

Objetivos específicos

- a) Analisar métodos diretos e iterativos de solução do problema de mínimos quadrados linear, quando o sistema formado por poucas épocas de observações GPS, tem indicativas de mal-condicionamento;
- b) Propor métodos de convergência global para a solução do problema de mínimos quadrados não-linear;
- c) Desenvolver métodos integrados para a resolução das ambigüidades GPS.

1.4 METODOLOGIA

Um dos objetivos dessa pesquisa é a verificação quanto à estabilidade dos métodos de solução do problema de mínimos quadrados linear. Sabe-se que em poucas épocas observadas, a geometria dos satélites quase não muda. A consequência é que as linhas da matriz Jacobiano são próximas de serem linearmente dependentes, conduzindo a problemas de mal-condicionamento. Neste sentido foram implementados cinco métodos, todos visando testar a estabilidade na solução e tempo computacional envolvido. Os métodos testados foram: Cholesky, decomposição QR, decomposição de valor singular (SVD), gradientes conjugados e gradientes conjugados com pré-condicionamento.

Para a solução do problema de mínimos quadrados não-linear são apresentadas duas estratégias de convergência global: uma baseada em região de confiança utilizando o compromisso de Levenberg-Marquardt e outra com algoritmo de *bactracking*, cuja direção de descida é definida a partir da decomposição de Cholesky modificada. Tal método foi denominado de Gauss-Newton modificado globalizado.

Os métodos propostos para a resolução das ambigüidades nesse trabalho foram dois; ambos classificados como métodos integrados. O primeiro, busca resolver as

ambigüidades por meio do problema de mínimos quadrados linear misto. Para isso, utilizam-se métodos de penalidades com injunções funcionais de desigualdade. Dentre as funções de penalidade, destaca-se a função de penalidade hiperbólica. Também, por causa da natureza inteira das ambigüidades, são utilizadas funções de penalidades sobre os não-inteiros (ver seção 4.4.1.3). O segundo, procura resolver as ambigüidades por meio de buscas direcionais, relacionando o espaço das ambigüidades e das coordenadas, aplicando-se para a direção definida, a solução do problema da função de ambigüidade (equação 1.7). Nesse método, também são utilizados conceitos de satélites primários e secundários para a validação das ambigüidades encontradas. As direções de buscas são definidas usando informações da matriz variância-covariância (ou da matriz cofatora) das ambigüidades reais estimadas.

1.5 ESTRUTURA DO TRABALHO

O trabalho está dividido em 5 capítulos incluindo o de conclusões. No primeiro capítulo, foram introduzidas as considerações gerais sobre as observações GPS, o conceito de ambigüidade, os principais problemas a serem tratados, os objetivos do trabalho e as metodologias empregadas. No segundo capítulo, relata-se os principais aspectos do GPS envolvidos no trabalho. No terceiro capítulo, aborda-se os conceitos e métodos relacionados com os problemas de mínimos quadrados (1.4), (1.5) e (1.6). No quarto capítulo, são apresentados alguns dos métodos de resolução de ambigüidades atualmente existentes e consagrados, nos quais fundamentam-se os métodos propostos. No quinto capítulo, são apresentadas as conclusões da pesquisa e algumas recomendações para futuras investigações sobre os métodos propostos.

2 FUNDAMENTOS DO GPS

O GPS, conforme mencionado na seção 1.1, começou a ser desenvolvido a partir de 1973 sob os auspícios do *U.S. Department of Defense*- DoD para suprir necessidades de aplicações militares de obtenção do posicionamento, tempo e velocidade. A implementação do sistema foi um sucesso, superando as expectativas até mesmo de seus idealizadores. Com a abertura à comunidade mundial, o GPS se tornou o sistema mais popular de navegação e posicionamento atualmente disponível.

A comunidade científica por sua vez, não ficou atrás e passou a explorar cada uma das potencialidades que o sistema oferece, a começar pelo monitoramento do deslocamento de placas tectônicas até a realização de sistemas de referência, como é o caso do Sistema de Referência Geocêntrico para a América do Sul – SIRGAS.

A descrição do GPS é costumeiramente dividida em três segmentos: o espacial, o de controle e o do usuário.

2.1 SEGMENTO ESPACIAL

2.1.1 Configuração Orbital e Satélites GPS

O sistema dispõe de 21 satélites em operação e mais 3 de reserva. Os satélites foram colocados em órbitas quase circulares em 6 planos orbitais, com uma inclinação de 55°. A altitude orbital é em torno de 20.200 km, correspondendo a 26.600 km para o semi-eixo maior. O período orbital é de 12 horas siderais (SEEBER, 1993, p. 211).

Segundo HOFMANN-WELLENHOF, LICHTENEGGER e COLLINS (1994, p.15), existem 5 classes ou tipos de satélites GPS:

Bloco I : satélites de desenvolvimento (*development satellites*);

Bloco II: satélites de produção (*production satellites*);

BlocoII A: satélites avançados de produção (*production advanced satellites*);

Bloco IIR: satélites de substituição (*replenishment satellites*);

Bloco IIF: satélites de continuidade (*follow on satellites*).

Os satélites do Bloco I foram lançados com 63° de inclinação, entre 22 de fevereiro de 1978 e 9 de outubro de 1985 num total de 11, sendo que em 1 deles ocorreu falha no lançamento. Pesavam 845 kg e não estavam habilitados com os códigos de segurança *Anti-spoofing* - AS e *Selective Availability* - SA. A frequência padrão era mantida por osciladores de quartzo.

Os satélites do bloco II foram lançados com 55° de inclinação, entre 14 de fevereiro de 1989 e 01 de outubro de 1990 num total de 9. Pesam cerca de 1.500 kg, e estão habilitados com o AS e o SA. Os osciladores de bordo são de césio e rubídio com estabilidade entre 10^{-12} e 10^{-13} segundos por dia.

Os satélites do bloco IIA foram lançados com 55° de inclinação, entre 26 de novembro de 1990 e 06 de novembro de 1997. Possuem ligeiras modificações em relação aos satélites do bloco II, têm capacidade de memória de 180 dias e alguns deles são equipados com refletores podendo ser rastreados com o sistema de medição a laser.

Os satélites do bloco IIR, em número de 20, deverão substituir os demais quando necessário. Possuem a capacidade de medir a distância entre eles (*cross link ranges*) e calcular as efemérides a bordo. Completamente operacionais, poderão ficar até 6 meses sem atualização do segmento de controle e ainda manterão a precisão das efemérides transmitidas (LEICK, 1995, p.62). A frequência padrão está planejada para ser mantida por oscilares de *Microwave Amplification by Stimulated Emission of Radiation* - MASER a hidrogênio com estabilidade de 10^{-14} a 10^{-15} segundos por dia (HOFMANN-WELLENHOF; LICHTENEGGER; COLLINS, 1994, p.18). Entretanto, os 5 primeiros satélites deste bloco estão operando com osciladores de rubídio (ver GPS WORLD SHOWCASE, 2000, p. 50).

Os satélites do bloco IIF deverão estar mais bem equipados do que os satélites anteriores, inclusive com sistema de navegação inercial (HOFMANN-WELLENHOF;

LICHTENEGGER; COLLINS,1994, p.18). A previsão de iniciar os lançamentos a partir de 2001 deve atrasar porque até o momento os satélites do bloco IIR não foram todos lançados (tab. 2.1).

A tab. 2.1 mostra a situação atual da constelação de satélites GPS. Na primeira coluna, aparece a seqüência de lançamento dos satélites e qual o bloco a que eles pertencem. Na segunda coluna, a identificação pelo número dos códigos pseudo-aleatórios (PRNs). Na terceira coluna, a data de lançamento. Na quarta coluna, a posição no plano orbital (os 6 planos orbitais são representados por A,B,C,D,E e F) e, na quinta coluna, a situação operacional atual do satélite.

TABELA 2.1 – SITUAÇÃO DA CONSTELAÇÃO DE SATÉLITES GPS (ATUALIZAÇÃO / 2001)

Continua

BLOCO SEQÜÊNCIA DE LANÇAMENTO (SV)	PRN	DATA DE LANÇAMENTO	POSICÃO NO PLANO ORBITAL	SITUAÇÃO ATUAL
BLOCO I				
01	04	22/02/78	-	Inativo 17/07/85
02	07	13/05/78	-	Inativo 16/07/81
03	06	06/10/78	-	Inativo 18/05/92
04	08	10/12/78	-	Inativo 14/10/89
05	05	09/02/80	-	Inativo 28/11/83
06	09	26/04/80	-	Inativo 06/03/91
07	***** FALHA NO LANÇAMENTO *****			
08	11	14/07/83	-	Inativo 04/05/91
09	13	13/06/84	C1	Inativo 20/06/94
10	12	08/09/84	A1	Inativo 18/11/95
11	03	09/10/85	C4	Inativo 13/04/94
BLOCO II				
01	14	14/02/89	E1	Ativo
02	02	10/06/89	B3	Ativo
03	16	18/08/89	E5	Inativo 13/10/00
04	19	21/10/89	A4	Ativo
05	17	11/12/89	D3	Ativo
06	18	24/01/90	F3	Inativo 18/08/00
07	20	26/03/90		Inativo 10/05/96
08	21	02/08/90	E2	Ativo
09	15	01/10/90	D2	Ativo
BLOCO IIA				
10	23	26/11/90	E4	Ativo
11	24	04/07/91	D1	Ativo
12	25	23/02/92	A2	Ativo
13	26	07/07/92	F2	Ativo
14	28	10/04/92	C5	Inativo 25/04/92

TABELA 2.1 – SITUAÇÃO DA CONSTELAÇÃO DE SATÉLITES GPS (ATUALIZAÇÃO / 2001)

BLOCO	PRN	DATA DE	POSIÇÃO	SITUAÇÃO ATUAL
SEQÜÊNCIA DE		LANÇAMENTO	NO	
LANÇAMENTO			PLANO	
(SV)			ORBITAL	
15	27	09/09/92	A3	Ativo
16	01	22/11/92	F1	Ativo
17	29	18/12/92	F4	Ativo
18	22	13/02/93	B1	Ativo
19	31	30/03/93	C3	Ativo
20	07	13/05/93	C4	Ativo
21	09	26/06/93	A1	Ativo
22	05	30/08/93	B4	Ativo
23	04	26/10/93	D4	Ativo
24	06	10/03/94	C1	Ativo
25	03	28/03/96	C2	Ativo
26	10	16/07/96	E3	Ativo
27	30	12/09/96	B2	Ativo
28	08	06/11/97	A5	Ativo
BLOCO IIR				
01	12	* * * FALHA NO LANÇAMENTO * * * * *		
02	13	23/07/97	F5	Ativo
03	11	07/10/99	D2	Ativo
04	20	11/05/00	E1	Ativo
05	28	16/07/00	B5	Ativo
06	14	10/11/00	F1	Ativo

FONTES: (1) CARVALHO, L. D. **Análise das técnicas GPS atuais para os posicionamentos estáticos e cinemáticos em bases curtas**. Curitiba, 1999, 139 f. Dissertação (Mestrado em Geodésia) – Setor de Ciências da Terra, Universidade Federal do Paraná.

(2) GPS WORLD SHOWCASE 2001. GPS Constellation, Published monthly by advanstar communications. v.11, n. 12, p. 50, 2000.

2.1.2 Sinais Transmitidos pelos Satélites

Os relógios atômicos a bordo dos satélites GPS produzem uma frequência fundamental na banda L de 10,23 MHz. A partir dela, estrategicamente, são derivadas coerentemente as ondas portadoras L_1 e L_2 , pela multiplicação eletrônica dos valores 154 e 120, respectivamente:

$$L_1 : f_1 = 1575,42 \text{ MHz} = 154 \times 10,23 \text{ MHz} \quad (2.1)$$

$$L_2 : f_2 = 1227,6 \text{ MHz} = 120 \times 10,23 \text{ MHz}. \quad (2.2)$$

A portadora L_1 é modulada pelos códigos pseudo-aleatórios (*pseudo random noise*- PRN) C/A e P, enquanto que a portadora L_2 é modulada apenas com o código P. De acordo com HOFMANN-WELLENHOF, LICHTENEGGER e COLLINS (1994, p.18) a omissão do código C/A na L_2 , é uma forma de controlar as informações transmitidas pelos satélites, degradando a precisão para usuários não militares.

Segundo SPILKER JUNIOR (1980, p.37-38), aos códigos C/A e P são adicionadas em módulo-2 as mensagens D com frequência de 50 bps. Assim, o sinal transmitido em fase e quadratura na portadora L_1 pelo satélite i , no instante t , tem a forma:

$$S_{L_1}^i(t) = A_p P^i(t) D^i(t) \cos(2\pi f_1 t + \phi) + A_c G^i(t) D^i(t) \sin(2\pi f_1 t + \phi), \quad (2.3)$$

onde: f_1 é a frequência conforme definida em (2.1); ϕ representa um pequeno ruído inicial da fase devido a uma deriva no oscilador; P e G representam os códigos (PRNs ± 1), P e C/A, respectivamente; A_p e A_c são constantes da amplitude dos códigos P e C/A, respectivamente.

Na portadora L_2 o sinal transmitido tem a forma:

$$S_{L_2}^i(t) = B_p P^i(t) D^i(t) \cos(2\pi f_2 t + \phi), \quad (2.4)$$

onde: B_p é amplitude do sinal; $P^i(t)$ é o código P do satélite i sincronizado com o código na L_1 .

2.1.3 Degradação da Acurácia por Meio de Técnicas de Segurança

O GPS, por se tratar de um sistema elaborado para fins militares, possui dispositivos para diminuir a acurácia do posicionamento e navegação de usuários não militares.

A primeira limitação denominada *Anti-spoofing* – AS é no acesso ao código P, o qual é modificado através de um código secreto W de encriptamento, com uma frequência de aproximadamente 500 KHz. O código P encriptado pelo código W,

passou a ser denominado de código Y ou P(Y). A técnica foi efetivada no seu completo potencial a partir de janeiro de 1994.

Sem o acesso direto ao código P, os receptores devem lançar mão de alguma técnica (ver seção 2.3.2) para obter as medidas de pseudodistância e fase da portadora na L_2 . As observáveis na portadora L_2 constituem um papel fundamental na correção dos erros devido aos efeitos da ionosfera.

A segunda limitação denominada de *Selective Availability*–SA utiliza dois meios para degradar a acurácia do posicionamento e navegação. Um, o processo ϵ , realiza manipulação das mensagens de navegação nas efemérides transmitidas, e o outro, o processo δ , falsifica o relógio do satélite introduzindo uma variação pseudo-aleatória em sua frequência.

O SA foi colocado em operação a partir de 25 de março de 1990, mas atualmente, o processo δ do SA está desligado. Isso aconteceu às 4 horas do Tempo Universal – TU do dia 2 de maio de 2000 (DIVIS, 2000, p. 16) ; (GPS WORLD, 2000, p. 20).

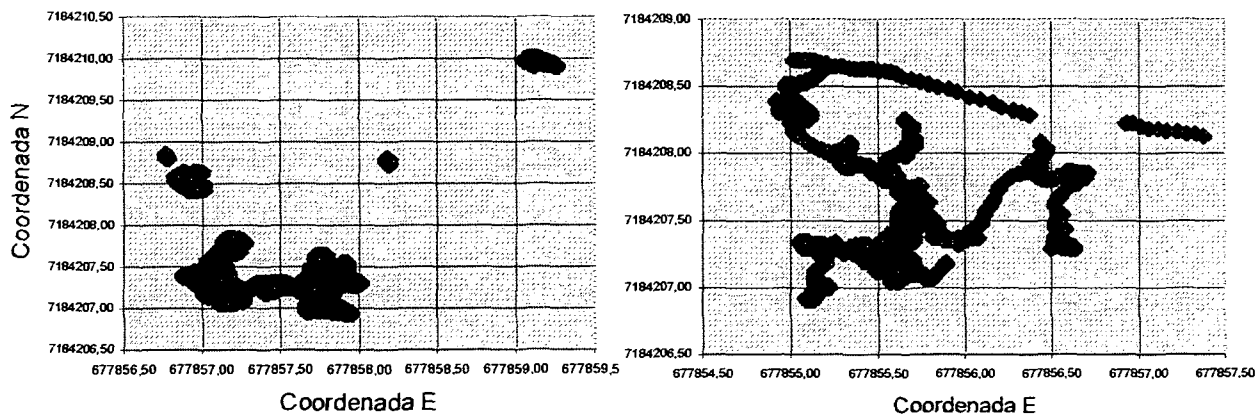
Experiências realizadas por MACHADO; MONICO e CAMARGO (2000, p. 38-43) com o satélite PRN 10 mostraram que o erro na posição orbital do satélite, antes e depois do desligamento do SA se manteve abaixo de 13,6 m (o desligamento em nada afetou), ao passo que as variações do erro do relógio praticamente desapareceram após o desligamento. Ainda, segundo essas mesmas experiências os autores mostraram que a precisão no posicionamento de um ponto, com a parte δ do SA desligado, aumentou cerca de 5 vezes em relação à situação anterior.

A fig. 2.1 ilustra duas experiências de posicionamento absoluto após o desligamento do SA. Os levantamentos foram realizados espaçados de 22 dias um do outro. A posição da estação foi diretamente calculada pelo receptor Ashtech Z-12 e listada no formato *National Maritime Electronics Association*–NMEA (ASHTech, 1994, p.157). As discrepâncias máximas na posição, entre as coordenadas UTM da estação RM3/SIRGAS e as coordenadas calculadas na projeção UTM/WGS-84 foram de 3,31 m (fig.2.1a) e 3,50m (fig. 2.1b).

FIGURA 2.1 – POSICIONAMENTO ABSOLUTO SEM O δ - SA

2.1(a) – Levantamento dia 18/11/2000.
 Erro máximo na posição 3,31 m.
 Tempo 2^h 53^m 11^s – taxa 1^s.

2.1(b) – Levantamento dia 27/10/2000.
 Erro máximo na posição 3,50 m.
 Tempo 2^h 46^m 42^s – taxa 2^s.



NOTA: Coordenadas na projeção UTM / WGS-84.
 Estação levantada RM3, SIRGAS.
 Receptor utilizado Ashtech Z-12 - formato NMEA \$GPGGA

2.2 SEGMENTO DE CONTROLE

De acordo com SEEGER (1993, p. 213) a principal tarefa do segmento de controle é:

- monitorar e controlar continuamente o sistema de satélites;
- determinar o sistema de tempo GPS;
- prever as efemérides dos satélites e o comportamento dos relógios dos satélites;
- atualizar periodicamente as mensagens de navegação para cada satélite.

Para desempenhar as tarefas acima, o segmento utiliza 5 estações de monitoramento: Hawaii, Kwajalein no Oceano Pacífico, Ilha de Ascensão no Oceano Atlântico, Diego Garcia no Oceano Índico e Colorado Spring nos Estados Unidos.

Cada uma das estações de monitoramento possui receptores de dupla frequência conectados a um oscilador externo de césio. As estações de monitoramento rastreiam dados de todos os satélites visíveis e transmitem para a estação mestra

(Colorado Spring). Os dados são processados, determinando-se as órbitas dos satélites (*broadcast ephemerides*) e as correções dos relógios para atualização das mensagens de navegação. As informações atualizadas são transmitidas para cada satélite por meio das antenas localizadas em Ilha de Ascension, Diego Garcia e Kwajalein.

Além das 5 estações mencionadas, outras estações da *National Imagery and Mapping Agency* – NIMA vêm sendo incorporadas ao segmento, são elas: Alaska, USNO, Ecuador, Argentina, England, Bahrain, South África, China, South Korea, Austrália, New Zealand (SEEBER, 2000).

2.3 SEGMENTO DO USUÁRIO

O seguimento do usuário consiste de todos os usuários; o militar e o civil, que com receptores apropriados, rastreiam código ou código e fase da portadora. Dentro desse segmento também são desenvolvidos os receptores, antenas, fontes de alimentação e osciladores. A descrição desses tópicos é encontrada em SEEBER (1993, p. 229 - 246) e WELLS et al. (1986, cap.7). De forma resumida, serão descritos nas seções subseqüentes (2.3.1) e (2.3.2), alguns dos tipos de receptores e técnicas *codeless*.

2.3.1 Tipos de Receptores

SEEBER (1993, p. 232 e 233) apresenta 4 classificações de receptores, sendo: quanto ao tipo de tecnologia utilizada no processamento do sinal; quanto aos tipos de observáveis produzidas; quanto à realização técnica dos canais; e quanto à comunidade usuária.

Com relação ao tipo de observável, pode-se ter:

1. código C/A;
2. código C/A + fase da portadora L_1 ;
3. código C/A + fase da portadora em L_1 e L_2 ;

4. código C/A e P + fase da portadora em L_1 e L_2 .

Receptores com código C/A

- Utilizam medidas de pseudodistâncias apenas com o código C/A. Os receptores típicos possuem vários canais independentes e fornecem a posição tri-dimensional (X,Y,Z), ou coordenadas elipsoidais (ϕ, λ, h), ou ainda coordenadas UTM (N,E).

Receptores com código C/A + fase da portadora L_1

- Utilizam medidas do código C/A e da fase da portadora L_1 . Receptores deste tipo estão ganhando cada vez mais espaço no mercado, à medida em que redes geodésicas locais de serviços de transmissão de correções diferenciais estão sendo implantadas.

Receptores com código C/A + fase da portadora L_1 e L_2

- Utilizam medidas do código C/A e da fase da portadora L_1 . O código C/A não é modulado na portadora L_2 . Os receptores possuem capacidade de obter a fase da portadora L_2 através de alguma técnica *codeless*. A qualidade do sinal-ruído neste caso é bem inferior ao obtido da portadora L_1 . Estes receptores podem ser usados para todos os tipos de levantamentos precisos, incluindo estático, cinemático e pseudocinemático¹.

Receptores com código C/A e P + fase da portadora em L_1 e L_2

- Receptores deste tipo usam o código P para sintonizar as portadoras L_1 e L_2 . Com o AS ativado, o código P é substituído pelo código P(Y) e a técnica de correlação do código P não pode ser aplicada. No entanto, este tipo de receptor pode operar em um modo *codeless* ou *quasi-codeless* proporcionando medidas de fase e pseudodistância na

¹ As técnicas de levantamento são amplamente discutidas e testadas em TIERRA CRIOLLO (1993) e CARVALHO (1999).

L_2 sem o conhecimento do código $P(Y)$. Se o receptor estiver equipado com o *Auxiliar Output Chip* - AOC terá acesso direto ao código $P(Y)$ com o AS ativado. Para isto o usuário deve ter autorização do DoD.

2.3.2 Técnicas *Codeless* e *Quasi-codeless*

Essas técnicas são empregadas nos receptores das classes 3 e 4 (seção 2.3.1), principalmente para obter as medidas de fase na portadora L_2 . As 4 técnicas usadas são: quadratura do sinal (*signal squaring*), correlação cruzada (*cross-correlation*), quadratura após a correlação do código (*code-correlating squaring*), e *Z tracking*.

2.3.2.1 Quadratura do sinal (*signal squaring*)

O sinal recebido do satélite é multiplicado por si mesmo e gera um segundo harmônico com 2 vezes a frequência do original (conseqüentemente o comprimento de onda é reduzido à metade do original). As mensagens de navegação e os códigos são perdidos (SEEBER, 1993, p. 235). Para ver isso, basta tomar um sinal do tipo (2.4) e lembrar que o código é uma sequência binária +1 e -1.

Exemplo:

$$y = P(t) \cos(ft). \quad (2.5)$$

Quadrando tem-se:

$$y^2 = P(t)^2 \cos^2(ft) = \frac{1}{2} [1 + \cos(f' t)], \quad (2.6)$$

$$\text{onde } f' = 2f \quad \Rightarrow \quad \lambda' = \frac{\lambda}{2}.$$

2.3.2.2 Correlação cruzada (*cross-correlation*)

A técnica é baseada no fato do código $P(Y)$ embora desconhecido, é idêntico nas duas portadoras, permitindo assim, a correlação cruzada de ambos os sinais L_1 e L_2 . O resultado deste processo é a diferença de pseudodistância, $P_{L2}(Y)-P_{L1}(Y)$, e a diferença da fase $\varphi_{L2} - \varphi_{L1}$. Combinando esses resultados com a pseudodistância no código C/A e fase da portadora L_1 , obtém-se valores para a pseudodistância e fase da portadora na L_2 , conforme equações (2.7) e (2.8) (LEICK, 1995,p.90); (HOFMANN-WELLENHOF; LICHTENEGGER; COLLINS,1994, p.85)

$$P_{L2} = P_{L1,C/A} + [P_{L2}(Y)-P_{L1}(Y)], \quad (2.7)$$

$$\varphi_{L2} = \varphi_{L1,C/A} + [\varphi_{L2} - \varphi_{L1}] . \quad (2.8)$$

2.3.2.3 Quadratura após a correlação do código (*code-correlating squaring*)

Essa técnica realiza uma correlação entre o código $P(Y)$ recebido e o código P gerado no receptor. Isto é possível porque o código P constitui a maior parte do código $P(Y)$ existindo uma similaridade entre eles, isto é, a influência do código W de criptografia é pequena, a frequência é aproximadamente 20 vezes menor do que a do código $P(Y)$. Depois da correlação um filtro passa-baixo é aplicado por *narrowing* e *bandwidth* e subsequentemente o sinal é quadrado, obtendo-se assim, meio comprimento de onda na portadora L_2 (HOFMANN-WELLENHOF; LICHTENEGGER; COLLINS, 1994, p.85).

2.3.2.4 Z Tracking

Essa é uma técnica *quasi-codeless*, também se baseia no princípio de que o código P tem maior contribuição no código recebido $P(Y)$.

É realizada uma correlação entre o código P gerado no receptor e os sinais em

L_1 e L_2 , a largura de banda (*bandwidth*) é reduzida para a do código W. O código W é estimado e removido a partir do sinal recebido permitindo que as réplicas do código gerado localmente sejam sintonizadas (*locked*) com os sinais do código P em L_1 e L_2 . O método não requer o conhecimento do código de criptografia, mas requer alguns aspectos de tempo. A técnica fornece as pseudodistâncias P_{L1} e P_{L2} e as fases da portadora ϕ_{L1} e ϕ_{L2} com o comprimento de onda completo (LEICK, 1995, p.91).

Essa técnica é também identificada por P-W Tracking (KRUEGER, 1996, p.40) e (MONICO, 1995, p.14).

O quadro 2.1 indica as principais características das 4 técnicas descritas anteriormente. Note que o valor de referência para o sinal ruído (SNR), é o da técnica *Z Tracking*.

QUADRO 2.1 – COMPARAÇÃO DAS TÉCNICAS USADAS PARA O AS

PARÂMETRO	QUADRATURA	CORRELAÇÃO CRUZADA	CORRELAÇÃO CÓDIGO + QUADRATURA	Z TRACKING
Código C/A	Não	Sim	Sim	Sim
Pseudodistância P(Y)	Não	$P(Y_2)-P(Y_1)$	$P(Y_1)$	$P(Y_1), P(Y_2)$
Comp. De onda L_2	Meio	Completo	Meio	Completo
SNR	-16 dB	-13 dB	-3 dB	0 dB

FONTE: Adaptado de LEICK, A. *GPS satellite surveying*. 2. ed. New York: John Wiley & Sons, Inc., 1995, p.91.

2.4 SISTEMA DE TEMPO GPS

A escala de tempo GPS é derivada a partir dos relógios das estações de controle, e sua origem é às 0^h do Tempo Universal Coordenado - TUC do dia 6 de janeiro de 1980 (domingo). A escala de tempo GPS é constante, ao passo que, o TUC, sofre correções periódicas para manter a relação (LEICK, 1995, p. 35):

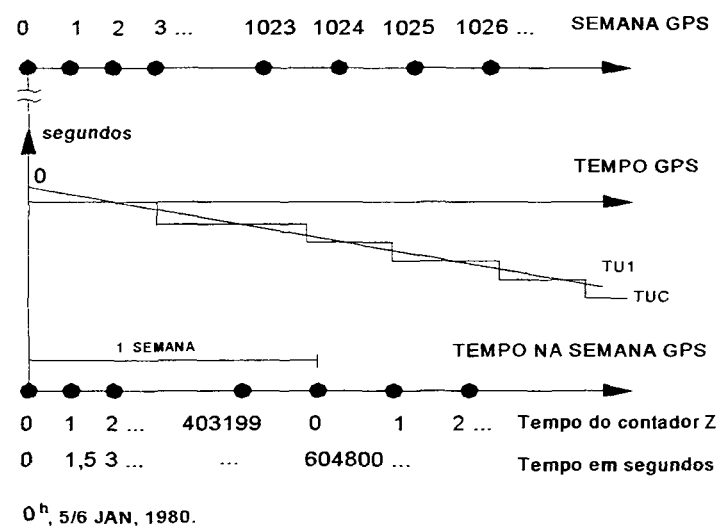
$$|TUC - TU_1| \leq 0,9 \text{ s} \quad (2.9)$$

onde: TU_1 é o Tempo Universal corrigido do movimento do pólo.

Manter a diferença mostrada na equação (2.9) é uma das responsabilidades do *International Earth Rotation Service – IERS*.

Faz parte do sistema de tempo GPS, a contagem da semana GPS, pois os parâmetros de navegação transmitidos pelos satélites têm como época de referência, a origem da semana. A numeração da semana foi inicialmente programada para ir de 0 a 1023, iniciando-se no dia 6 de janeiro de 1980, até aproximadamente mais uns 20 anos. Hoje, já tendo passadas as 1023 semanas, por questões de operacionalidade, a contagem para alguns serviços como a do *International GPS Service* – IGS e a do *National Geodetic Survey* – NGS, prosseguem normalmente. Na fig.2.2, ilustra-se esquematicamente a relação entre o tempo GPS, TUC e TU_1 .

FIGURA 2.2 – SISTEMA DE TEMPO GPS



FONTE: Adaptada de ANDRADE, J. B. de. **NAVSTAR-GPS**. Curitiba, 1988. 63 f. Apostila - Curso de Pós-Graduação em Ciências Geodésicas, Universidade Federal do Paraná, p.12.

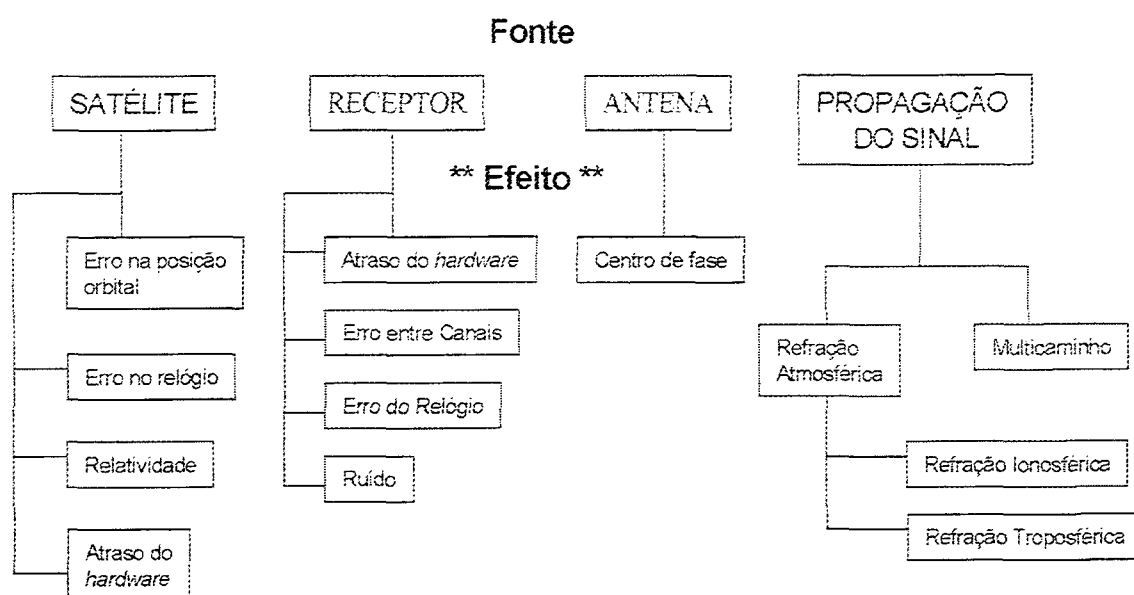
2.5 PRINCIPAIS ERROS DO GPS

Do ponto de vista estrito, qualquer sistema que envolva medidas está sujeito a algum tipo de erro, seja sistemático ou aleatório. No GPS não é diferente, há fontes de erros nos satélites, no receptor, nas antenas, na estação base no caso do posicionamento relativo e na propagação do sinal. Além dessas fontes, existem os erros de operação em campo em relação aos procedimentos, por exemplo: medir corretamente a altura da antena, orientá-la em azimuth quando necessário, centrar com

a marca de referência da estação, nivelar a base, etc.

Não constitui objeto desse trabalho uma completa revisão sobre todos os tipos de erros que afetam a alta precisão do posicionamento GPS. Porém, é importante revisar alguns. O diagrama ilustrado na fig. 2.3 mostra alguns destes erros, relacionando a fonte e o efeito.

FIGURA 2.3 – ERROS DO GPS



2.5.1 Erro na Posição Orbital

Para o usuário que não está realizando pesquisa na determinação orbital do GPS, as posições do satélite em um determinado instante de transmissão devem ser consideradas conhecidas, isto é, não vão entrar como parâmetro no processamento dos dados.

Se o processamento a ser realizado é em tempo real ou logo após o término do levantamento, a opção é determinar a posição dos satélites utilizando as efemérides de navegação transmitidas. Alguns detalhes do algoritmo para o cálculo das coordenadas

dos satélites usando as efemérides transmitidas, conforme apresentado em LEICK (1995, p. 75), foram discutidos e esclarecidos em GARNÉS e FERREIRA (1999).

Se o processamento puder ser executado vários dias ou várias horas depois das observações terem sido realizadas, existe a opção de se utilizar órbitas mais precisas produzidas por algumas das redes oficiais.

O *International GPS Service* – IGS começou as campanhas de testes em 21 de junho de 1992, utilizando principalmente observações da fase da portadora em uma rede densa de estações em todo o mundo. A partir de 30 de Junho de 1996 o IGS passou a disponibilizar as órbitas rápidas com 24 h de atraso e as órbitas finais com 11 dias de atraso (ROTHACHER et al., 1996, p. 103). Hoje, no entanto, o Serviço disponibiliza órbitas preditas ultra-rápidas (duas vezes por dia), órbitas rápidas (com 17 horas de atraso) e órbitas finais (com aproximadamente 13 dias de atraso). A tab. 2.2 ilustra um comparativo entre os tipos de órbitas.

As órbitas IGS fornecem as coordenadas geocêntricas dos satélites de 15 em 15 minutos com referência ao *IERS Terrestrial Reference Frame* – ITRF, no formato denominado SP3 (BEUTLER et al., 1998, p.100).

TABELA 2.2 – ÓRBITAS DO GPS

TIPO DE ÓRBITA	ACURÁCIA	TEMPO PARA DISPONIBILIDADE	LOCAL DA DISPONIBILIDADE
Órbita transmitida	~ 2,60 m	Tempo real	Mensagem transmitida
Órbita predita (ultra rápida)	~ 0,25 m	Tempo real	Centro de dados IGS
Órbita rápida IGS	0,05 m	Depois de 17 h	Centro de dados IGS
Órbita final IGS	< 0,05 m	Depois de 13 dias	Centro de dados IGS

FONTE: IGS Products. Disponível em: <<http://igscb.jpl.nasa.gov/components/prods.html>>
Acesso em: 1 jun. 2001.

Existem vários outros Serviços que disponibilizam órbitas precisas, por exemplo, o *National Geodetic Service* – NGS e a *National Oceanic Atmospheric Administration* – NOAA, são responsáveis pelo fornecimento das efemérides precisas disponibilizadas no site do *United States Coast Guard Navigation Center*, no endereço: <<http://www.navcen.uscg.gov/gps/precise/default.htm>> Acessado em: 1 ago. 2001.

Para o posicionamento relativo foram desenvolvidas regras práticas relacionando a magnitude do erro do satélite com o erro da base. BAUERSIMA², citado por BEUTLER et al. (1998, p.104), em 1983 fornece a seguinte regra:

$$\frac{\|\mathbf{db}\|}{\|\mathbf{b}\|} = \frac{\|\mathbf{dr}\|}{\|\mathbf{r}\|} \quad (2.10)$$

e, ZIELINSKI³, citado por BEUTLER et al. (1998, p.104), em 1988 baseando-se em métodos estatísticos apresenta a seguinte modificação:

$$\frac{\|\mathbf{db}\|}{\|\mathbf{b}\|} = \frac{\|\mathbf{dr}\|}{\|\mathbf{k r}\|}, \text{ com } 4 < k < 10. \quad (2.11)$$

Segundo BEUTLER et al.(1998, p.104), a equação (2.10) é somente mais adequada para verificar a componente da altitude da linha de base, enquanto que a equação (2.11) é mais justa para a linha de base em si, isto é, para as componentes: latitude (ϕ), longitude (λ) e altitude (h).

A tab. 2.3, mostra quanto um erro $\|\mathbf{dr}\| = 100$ m nas posições dos satélites se propaga nas linhas de base do posicionamento relativo de até 100 km.

TAB. 2.3 – INFLUENCIA DO ERRO ORBITAL NO POSICIONAMENTO RELATIVO

Comp. Da linha de base $\ \mathbf{b}\ $	Erro da linha de base $\ \mathbf{db}\ $
1 km	0,6 mm
10 km	5,7 mm
20 km	11,4 mm
40 km	22,9 mm
60 km	34,3 mm
80 km	45,7 mm
100 km	57,1 mm

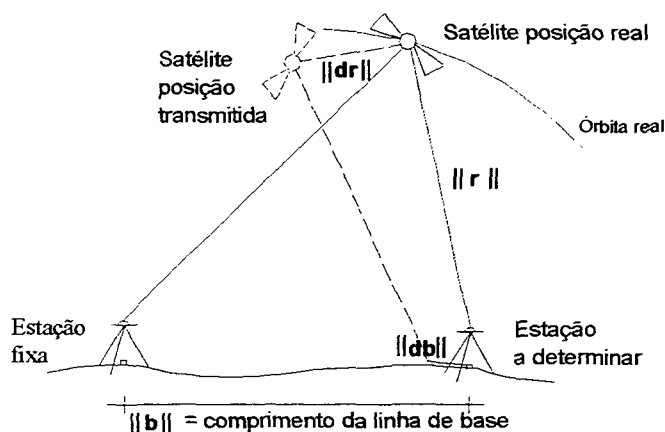
NOTA: $\|\mathbf{dr}\|=100\text{m}$, $r = 25.000$ km e $k=7$ na equação (2.11)

² BAUERSIMA, I. Navstar/Global Positioning System (GPS) (II) “Mitteilung no. 10 der Satellitenbeobachtungsstation Zimmerwald, Druckerei der Universität Bern, 1983.

³ ZIELINSKI, J. B. “Covariances in 3D network resulting from orbital errors”. Proceedings of the International GPS-Workshop, Darmstadt, Abril 10-13, Published in Lecture Notes in Earth Sciences, GPS-Techniques Applied to Geodesy and Surveying, Springer Verlag, Berlin, p. 504-514, 1988.

A fig. 2.4 ilustra os elementos envolvidos na equação (2.10) (ou equação (2.11)).

FIGURA 2.4 – ERRO NA POSIÇÃO ORBITAL X ERRO NA LINHA DE BASE



2.5.2 Erro do Relógio do Satélite

As estações de controle responsáveis pela manutenção do tempo GPS constantemente monitoram o rendimento dos relógios dos satélites. A diferença desses relógios em relação ao padrão é mantida dentro de 1 ms (milissegundo). A correção é transmitida ao segmento do usuário através das mensagens de navegação pelos coeficientes de um polinômio do segundo grau (WELLS et al., 1986, p.9.2):

$$dt^s = a_0 + a_1(t-t_{oc}) + a_2(t-t_{oc})^2, \quad (2.12)$$

onde:

t_{oc} : época de referência dos relógios dos satélites;

a_0, a_1, a_2 : coeficientes do polinômio, transmitidos;

dt^s : estado (erro) do relógio do satélite para o instante t .

A sincronização entre os relógios dos satélites e o tempo GPS é mantida dentro de 20 ns (nanosegundo) utilizando a correção das efemérides transmitidas (WELLS et al., 1986, p. 9.2).

Na presença do δ -SA a equação que descreve a frequência no relógio do satélite é (SEEBER, 1993, p.301)

$$f_i(t) = f_0 + \Delta f + \dot{f}(t-t_{oc}) + \delta f_{SA}(t) + \tilde{f}(t), \quad (2.13)$$

onde:

f_0 : frequência fundamental 13,23 MHz;

Δf : diferença sistemática da frequência;

\dot{f} : deriva da frequência;

δf_{SA} : erro na frequência causado pelo SA; e

\tilde{f} : erro aleatório da frequência.

Segundo o mesmo autor, o único termo desconhecido na equação (2.13) é o termo devido ao efeito do SA

$$\delta t_{SA} = T \int_{t_0}^t \delta f_{SA}(t) dt, \quad (2.14)$$

sendo T o período. Esse efeito pode ser removido utilizando-se o processo relativo quando as observações são simultâneas.

2.5.3 Relatividade

Devido à diferença entre o potencial gravitacional do satélite em órbita e o potencial gravitacional na superfície física da Terra, o relógio do satélite adianta – efeito da relatividade geral. Devido à velocidade do satélite, o relógio atrasa – efeito da relatividade especial (GEMAEL, 1991, p. 8.13).

Os erros por esses efeitos, em parte, são minimizados reduzindo a frequência nominal dos relógios de bordo (10,23 MHz) em 0,00455 Hz antes do lançamento (ANDRADE, 1988, p.14). Mesmo assim, resta uma pequena componente constante devido à diferença da altitude do satélite com a sua altitude nominal e, ainda, uma outra componente periódica. O efeito constante é absorvido pelo coeficiente a_1 do polinômio na equação (2.12), e o efeito periódico pode ser corrigido por (SEEBER, 1993, p. 292):

$$\Delta t_R = -4,443 \cdot 10^{-10} e \sqrt{A} \sin(E), \quad (2.15)$$

onde:

Δt_R : correção do efeito relativístico periódico (s);

e : excentricidade da órbita;

A : semi-eixo maior da órbita (m);

E : anomalia excêntrica.

2.5.4 Atraso do *Hardware*

O atraso de *hardware* ocorre tanto no satélite quanto no receptor. É o atraso que o sinal experimenta ao percorrer os componentes eletrônicos do equipamento. No satélite corresponde ao tempo entre a geração do sinal e a transmissão a partir da antena. No receptor corresponde ao tempo de percurso da recepção na antena até o “correlator” de sinal no receptor (TEUNISSEN; KLEUSBERG, 1998, p. 190).

2.5.5 Erro entre Canais (*Interchannel Biases*)

Em receptores que possuem vários canais, cada canal é dedicado à recepção do sinal de um satélite. Se os canais não trabalharem totalmente sincronizados, haverá uma diferença sistemática entre eles. Para corrigir eventuais diferenças entre diferentes canais, o receptor realiza uma calibração intercanais no início de cada levantamento.

Isto envolve cada canal rastrear simultaneamente um satélite e determinar o desvio em relação a um canal de referência. Todas as medidas subsequentes são ajustadas com esta diferença (MONICO,1995, p.28).

2.5.6 Erro do Relógio do Receptor

Os receptores geralmente são equipados com relógios de cristal de quartzo de alta qualidade (WELLS et al.,1986, p. 9.2). Contudo, o sincronismo entre o tempo do receptor e o tempo GPS não é perfeitamente mantido. Essa falta de sincronia ou estado (erro) do relógio do receptor em relação ao tempo GPS, dt_R , pode ser estimado em cada época, desde que 4 ou mais satélites sejam rastreados simultaneamente. Ou pode ser modelado ao longo da seção de rastreio por um polinômio, por exemplo:

$$dt_R = b_0 + b_1 (t-t_0) + b_2 (t-t_0)^2 , \quad (2.16)$$

onde: t_0 pode ser a primeira época observada; e b_0 , b_1 e b_2 são os coeficientes a estimar no ajustamento.

2.5.7 Ruído no Receptor

O ruído está relacionado com a resolução das observações, isto é, a limitação dos receptores eletrônicos na realização das medidas. Tem caráter aleatório e geralmente está associado a 1% do comprimento de onda (WELLS et al., 1986, p.9.14). Na tab.1.1 (ver seção 1.1), estão indicadas as magnitudes dos ruídos das observáveis dos códigos e fases da portadoras.

2.5.8 Centro de Fase

O posicionamento de um ponto GPS é realizado com base nas observações a partir do sinal recebido na antena. A localização de referência no registro dessas

informações é o centro de fase da antena. Desta forma, no posicionamento relativo, uma linha de base, em princípio, fica determinada pelo centro de fase das duas antenas utilizadas. Ocorre, porém, que em geral o centro de fase e o centro geométrico (mecânico) das antenas GPS não são coincidentes. Além disso, o centro de fase para a portadora L_1 é diferente do centro de fase para a portadora L_2 e, conseqüentemente, não será o mesmo para uma observável formada a partir da combinação linear delas.

Outra questão a ser considerada em se tratando do centro de fase das antenas é que este não é um ponto fixo na antena, pois, varia com a direção do sinal que chega, isto é, depende do azimute e elevação do satélite emissor.

A não consideração da variação do centro de fase faz com que a linha de base fique definida em relação aos centros médios de fases das duas antenas, caracterizados pela média ponderada de todos os centros de fase individuais de cada época observada (MADER, 1999, p.1). Assim, em posicionamentos que exigem alta acurácia (erros de poucos milímetros) é necessário corrigir a diferença (*offset*) entre o centro médio de fase nominal e o centro geométrico da antena, bem como, a variação dos centros de fase individuais a cada época observada em relação ao centro médio de fase nominal.

A determinação do centro médio de fase de uma antena e sua variação é feita mediante a calibração da antena. Apesar de cada antena fabricada ter seus próprios valores de calibração, o que se pode fazer quando a antena não foi calibrada, é usar os valores do seu modelo de fabricação. Para isso, o *National Geodetic Survey-NGS* já, há algum tempo, realiza esse tipo de serviço e fornece os valores de calibração: *offset* médio e as variações do centro de fase conforme elevação do satélite.

A tab. 2.4 ilustra a forma e disposição dos elementos de calibração do NGS, e a tab. 2.5, mostra um exemplo para os elementos da tab. 2.4, fornecendo a magnitude do centro de fase em relação à superfície de referência da antena (ver fig. 2.5). A parte inferior da tab. 2.5 indica os desvios padrão obtidos de 4 medidas no processo de calibração da antena. No exemplo ilustrado (tab. 2.5), estrategicamente utilizou-se o modelo de antena *Trimble Choke Ring*.

TABELA 2.4 – ELEMENTOS DE CALIBRAÇÃO DA ANTENA

ANTENA ID	DESCRIÇÃO									FONTE DE DADOS (# DE TESTES) ANO/MÊS/DIA
[Norte]	[Leste]	[Altura]								L ₁ <i>offset</i> (mm)
[90]	[85]	[80]	[75]	[70]	[65]	[60]	[55]	[50]	[45]	L ₁ fase na
[40]	[35]	[30]	[25]	[20]	[15]	[10]	[5]	[0]		elevação (mm)
[Norte]	[Leste]	[Altura]								L ₂ <i>offset</i> (mm)
[90]	[85]	[80]	[75]	[70]	[65]	[60]	[55]	[50]	[45]	L ₂ fase na
[40]	[35]	[30]	[25]	[20]	[15]	[10]	[5]	[0]		elevação (mm)

Fonte: MADER, G.L., GPS Antena Calibration at the National Geodetic Survey. National Geodetic Survey, NOS, NOAA, Silver Spring, MD. Disponível em: < <http://www.ngs.noaa.gov/ANTCAL/Files/summary.html> > Acesso em: 1 jun. 2001.

TABELA 2.5 – ELEMENTOS DE CALIBRAÇÃO DO CENTRO DE FASE

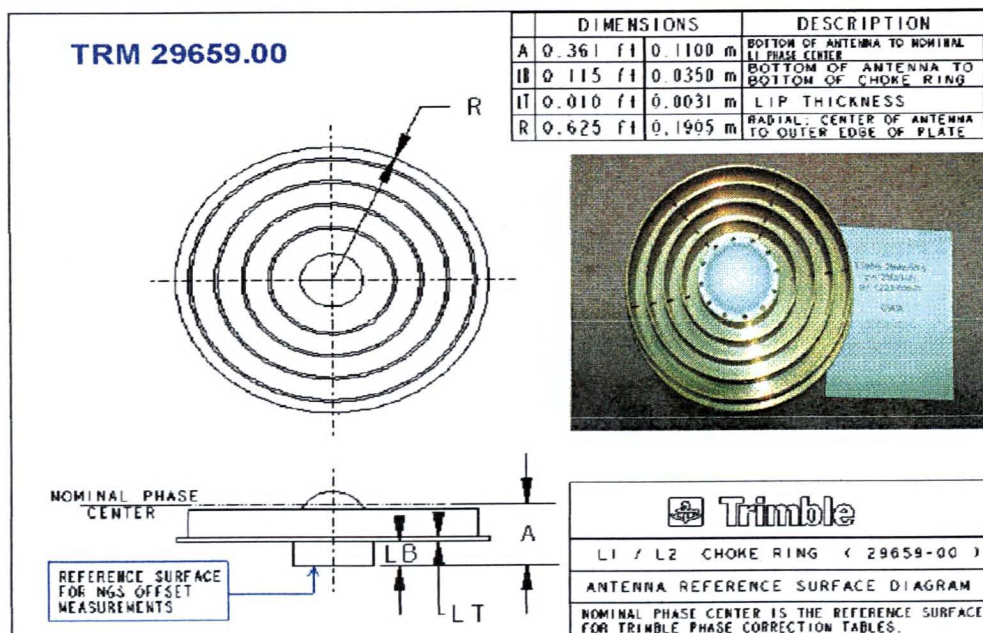
TRM29659.00	TRIMBLE CHOKE RING ANTENNA									NGS(4)97/10/27
1.2	.5	109.8								
.0	.3	.5	.5	.5	.5	.5	.5	.4	.4	.4
.4	.4	.4	.4	.5	.5	.4	.0	.0		
1.2	.6	128.0								
.0	.0	.0	.0	.0	.0	.1	.1	.0	.0	
-.1	-.1	-.1	.1	.1	.0	.2	.4	.0	.0	
RMS mm (1 sigma)	4 MEASUREMENTS									
.1	.2	.2								
.0	.3	.4	.5	.5	.5	.4	.4	.4	.4	.4
.4	.4	.4	.4	.4	.3	.2	.0	.0		
.2	.5	.2								
.0	.1	.2	.2	.2	.2	.2	.2	.2	.2	.2
.2	.2	.3	.3	.2	.2	.1	.0	.0		

FONTE : MADER, G.L., GPS Antena Calibration. Disponível em: < http://www.ngs.noaa.gov/ANTCAL/Models/TRM29659_00.shtml > Acesso em: 1 jun. 2001.

Uma das estações utilizadas nos testes desse trabalho e escolhida para melhor ilustrar essa seção, é a estação denominada PARA. Tal estação faz parte da realização do SIRGAS (IBGE, 1997, p.59) e constitui uma das estações brasileiras de monitoramento contínuo do Instituto Brasileiro de Geografia e Estatística - IBGE. Está localizada na área do Centro Politécnico, na Universidade Federal do Paraná - UFPR. Nesta estação está instalada uma antena Trimble Choke Ring.

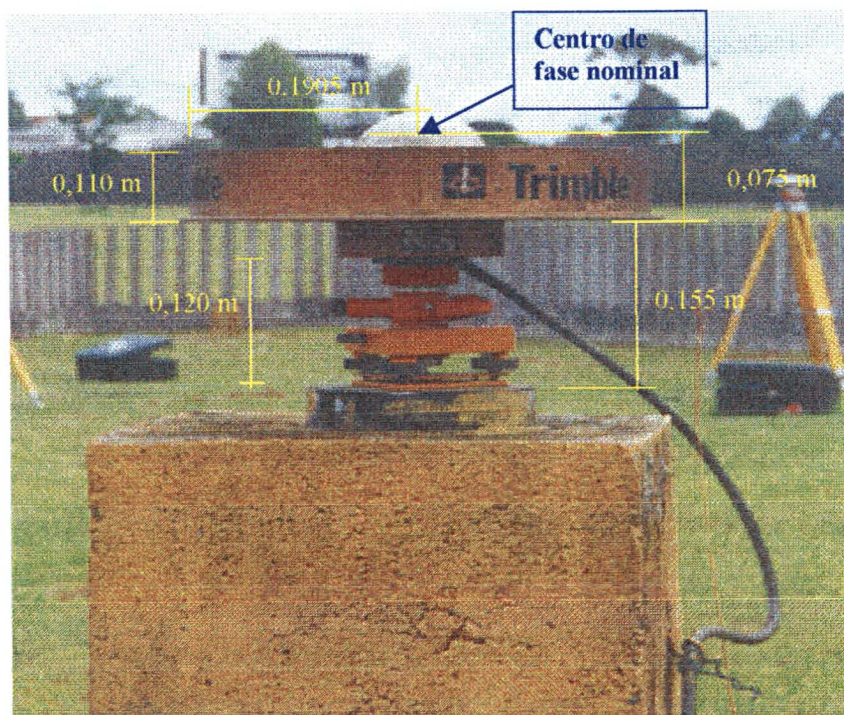
A tab.2.5 acima e as figs. 2.5 e 2.6, esclarecem sobre a importância em se considerar e corrigir o deslocamento (*offset*) do centro de fase das antenas nas aplicações de posicionamento preciso GPS.

FIGURA 2.5 – DIAGRAMA DE REFERÊNCIA DA ANTENA TRIMBLE CHOKE RING



FONTE: MADER, G.L. GPS Antena Calibration. **National Geodetic Survey, NOS, NOAA**. Disponível em: <http://www.ngs.noaa.gov/ANTCAL/Models/Diagrams/TRM29659_00.gif> Acesso em: 1 jun. 2001.

FIGURA 2.6 – CARACTERIZAÇÃO DO CENTRO DE FASE NA ESTAÇÃO PARA

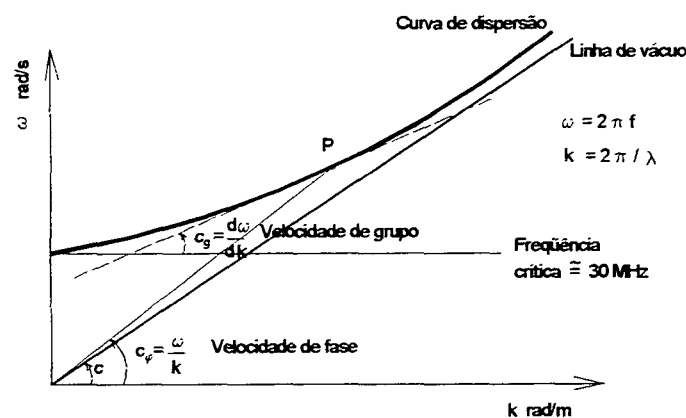


Quando não se têm os elementos de calibração das antenas (centros de fase e suas variações) a serem utilizadas em uma determinada campanha, recomenda-se o uso, no processo relativo, de um mesmo modelo de antena e uma correta orientação azimutal. Com isso, os erros do centro de fase deverão ser minimizados.

2.5.9 Refração Ionosférica

Para o GPS, a ionosfera geralmente compreende a camada da atmosfera entre aproximadamente 50 km a 1.000 km de altitude. Nesta região a radiação ultravioleta do Sol ioniza moléculas de gás que liberam elétrons livres. Os sinais do GPS, bem como, qualquer sinal eletromagnético ao se propagar pelo meio ionizado são afetados pelas características dispersivas do meio (WELLS et al., 1986, p.9.5)

FIGURA 2.7 – DISPERSÃO NA IONOSFERA



FONTE: Adaptado a partir de WELLS, D.; BECK, N.; DELIKARAOGLU, D.; KLEUSBERG, A.; KRAKIWSKY, E. J.; LACHAPELLE, G.; LANGLEY, R. B.; NAKIBOGLU, M. SCHWARZ, K. P.; TRANQUILLA, J. M.; VANÍCEK, P. **Guide to GPS positioning**. Canadá: Canadian GPS Associates, Fredericton, N. B., 1986, p. 9.05; e LEICK, A. **GPS satellite surveying**. 2. ed. New York: John Wiley & Sons, Inc., 1995, p. 293

Na curva de dispersão da ionosfera (fig. 2.7), a frequência angular (ω) é função do número de onda (k). A velocidade de fase (c_ϕ) de um ponto na curva é dada pela inclinação da reta que une este ponto à origem

$$c_{\varphi} = \frac{\omega}{k} = f \lambda . \quad (2.17)$$

Por sua vez, a velocidade de grupo (c_g) é dada pela inclinação da reta tangente no ponto

$$c_g = \frac{d\omega}{dk} . \quad (2.18)$$

No sinal GPS há que se considerar ambas as velocidades. A velocidade de grupo, quando são realizadas observações sobre os códigos e a velocidade de fase, quando são realizadas observações da fase da portadora.

Introduzindo os índices de refração de grupo (n_g) e de fase (n_{φ}), têm-se as seguintes relações:

$$n_{\varphi} c_{\varphi} = c; \quad (2.19)$$

$$n_g c_g = c, \quad (2.20)$$

onde: c é a velocidade da luz no vácuo.

A equação que relaciona os índices de fase e de grupo é (LEICK, 1995, p.293):

$$\begin{aligned} n_g &= \frac{c}{g} = c \frac{dk}{d\omega} = \frac{d}{d\omega} (n_{\varphi} \omega) \\ &= n_{\varphi} + f \frac{dn_{\varphi}}{df} . \end{aligned} \quad (2.21)$$

De acordo com SEEGER (1993, p.44) o índice de refração da fase pode ser desenvolvido como uma série de potência:

$$n_{\varphi} = 1 + \frac{c_2}{f^2} + \frac{c_3}{f^3} + \frac{c_4}{f^4} + \dots, \quad (2.22)$$

onde: os coeficientes c_2, c_3, \dots , não dependem da frequência, mas sim da densidade de

elétons Ne (elétons/m³), ao longo do percurso de propagação do sinal.

Utilizando a aproximação

$$n_{\varphi} = 1 + \frac{c_2}{f^2}, \quad (2.23)$$

o índice de refração estará cerca de 92% representado (ANDRADE, 1988, p.29).

Diferenciando a equação (2.23), tem-se:

$$dn_{\varphi} = -\frac{2c_2}{f^3} df \quad \Rightarrow \quad \frac{dn_{\varphi}}{df} = -\frac{2c_2}{f^3}. \quad (2.24)$$

Levando as equações (2.23) e (2.24) em (2.21) resulta:

$$n_g = 1 + \frac{c_2}{f^2}. \quad (2.25)$$

Uma estimativa para o valor de c_2 é (HOFMANN-WELLENHOF; LICHTENEGGER; COLLINS, 1994, p.101):

$$c_2 = -40,28 \text{ Ne Hz}^2. \quad (2.26)$$

Substituindo a (2.26) nas equações (2.23) e (2.25), tem-se, respectivamente:

$$n_{\varphi} = 1 - \frac{40,28 \text{ Ne}}{f^2}, \quad (2.27)$$

$$n_g = 1 + \frac{40,28 \text{ Ne}}{f^2}. \quad (2.28)$$

Como Ne é uma quantidade sempre positiva implica em $n_g > n_{\varphi}$ e conseqüentemente $c_{\varphi} > c_g$. Isto significa que a fase da portadora adianta da mesma quantidade em que o código atrasa ao atravessar a ionosfera.

A magnitude da refração ionosférica em unidades de comprimento é calculada por (LEICK, 1995, p.296):

$$d_{ion} = \frac{40,3}{f^2} \int_{\text{caminho}} Ne \, ds \quad (2.29)$$

A integral é definida como o conteúdo total de elétrons (*total electron content*) –TEC em uma coluna que se estende do receptor ao satélite com área de seção transversal igual a 1 m^2

$$\text{TEC} = \int_{\text{caminho}} Ne \, ds . \quad (2.30)$$

Com a equação (2.30), a equação (2.29) pode ser reescrita como:

$$d_{ion} = \frac{40,3}{f^2} \text{TEC} . \quad (2.31)$$

Segundo WELLS et al.,(1986,p.9.06) o valor extremo para d_{ion} no zênite chega a aproximadamente 50 m e no horizonte a aproximadamente a 150 m.

Se o receptor possui a capacidade de rastrear dados nas duas frequências, a refração pode ser corrigida em seu termo de primeira ordem e o erro resultante será de poucos centímetros (SEEBER, 1993, p.304).

Considerando as equações de observação das pseudodistâncias numa forma bastante simplificada (ver equação 2.62), tem-se:

$$P_1 = \rho + d_{ion,L1} , \quad (2.32)$$

$$P_2 = \rho + d_{ion,L2} , \quad (2.33)$$

onde:

P_1, P_2 : pseudodistâncias nas portadoras L_1 e L_2 ;

ρ : distância geométrica receptor/satélite;

$d_{ion,L1}$, $d_{ion,L2}$: refração ionosférica para a L_1 e L_2 , respectivamente.

Subtraindo a equação (2.33) da (2.32) e substituindo a (2.31) tem-se:

$$P_1 - P_2 = 40,3 \text{ TEC} \left(\frac{1}{f_1^2} - \frac{1}{f_2^2} \right). \quad (2.34)$$

Dividindo e multiplicando o segundo membro da equação (2.34) por f_1^2 e lembrando da (2.31), obtém-se a expressão para calcular a refração ionosférica para a L_1 .

$$d_{ion,L1} = \frac{(P_1 - P_2)f_2^2}{(f_2^2 - f_1^2)}. \quad (2.35)$$

Observação: com exceção do multicaminho que depende da frequência, os demais erros são comuns a ambas as frequências e com a diferença conforme equação (2.34), se cancelam. A refração, equação (2.35), deve ser calculada com ausência de multicaminho (ver multicaminho na seção 2.5.11).

Se o receptor estiver habilitado para receber dados de apenas uma frequência, a refração ionosférica poderá ser estimada por algum modelo, por exemplo, o modelo de KLOBUCHAR⁴, citado por LEICK (1995,p.301), utilizando os parâmetros das efemérides transmitidas.

O modelo de Klobuchar corrige cerca de 50 a 60% da refração ionosférica. Detalhes a respeito desse modelo são encontrados em LEICK (1995, p.302), CAMARGO (1999, p.92-96), e HOFMANN-WELLENHOF, LICHTENEGGER e COLLINS, (1994, p.105-106). Outros modelos para receptores de uma frequência são também encontrados em LEICK (1995) e CAMARGO (1999).

A outra opção para lidar com a refração ionosférica em receptores de uma frequência, é formar a combinação linear de observáveis entre receptores, como é o caso da simples diferença (ver seção 2.6.3.1). Segundo WELLS et. al., (1986, p.9.06) o residual típico da refração ionosférica resultante da simples diferença é 1 ppm.

⁴ KLOBUCHAR, J.A. "Ionospheric time-delay algorithm for single-frequency GPS users". IEEE Transactions on Aerospace and Electronic Systems, AES- 23(3),1987, p.325-331.

2.5.10 Refração Troposférica

A camada inferior da atmosfera com até aproximadamente um pouco mais de 50 km, constituída pela troposfera, tropopause, estratosfera e mesosfera (ver camadas da atmosfera em LEICK(1995, p.288) e LANGLEY(1998, p. 126)) possui átomos neutros e moléculas que afetam a propagação do sinal GPS. Esse meio é não dispersivo para sinais com frequência de até 30 GHz. Desse modo, tanto as medidas da pseudodistância com o código, como as medidas de fase da portadora, sofrem o mesmo efeito. Esse efeito, que retarda a propagação do sinal é conhecido por refração troposférica ou atraso troposférico.

A refração troposférica, conceitualmente, pode ser definida como a integral ao longo do caminho de propagação do sinal na camada neutra da atmosfera:

$$d_{trop} = \int_{\text{caminho}} (n - 1) ds, \quad (2.36)$$

onde: n é o índice da refração troposférica.

No desenvolvimento dos modelos da refração troposférica, costuma-se utilizar em vez do índice de refração (n), a refratividade:

$$N = (n-1) \cdot 10^6. \quad (2.37)$$

A refratividade pode ser escrita como a soma de duas componentes, uma seca e outra úmida:

$$N = N_d + N_w, \quad (2.38)$$

onde:

N_d : refratividade para o ar seco;

N_w : refratividade para o ar úmido.

Existem alguns modelos para calcular a refratividade em função das condições meteorológicas do lugar, um deles é o de Smith-Weintraub, com acurácia de 0,5% para frequências abaixo de 30GHz (LANGLEY,1998,p.128):

$$N = 77,6 \frac{P}{T} + 3,73 \cdot 10^5 \frac{e}{T^2}, \quad (2.39)$$

com:

$$N_d = 77,6 \frac{P}{T};$$

$$N_w = 3,73 \cdot 10^5 \frac{e}{T^2},$$

onde:

P : pressão atmosférica em mb (1 b = 100.000 Pa; 1 atm = 1.013,25 HPa);

T : temperatura Kelvin, K (K = 273,15 + °C);

e : pressão parcial do vapor d'água em mb.

A refração troposférica total é calculada por:

$$\begin{aligned} d_{trop} &= 10^{-6} \int_{caminho} N_d ds + 10^{-6} \int_{caminho} N_w ds \\ &= d_{trop,d} + d_{trop,w}, \end{aligned} \quad (2.40)$$

com a componente seca representando cerca de 90% do total.

Muitos modelos foram desenvolvidos para calcular a refração troposférica. Um bastante citado é o de Hopfield apresentado a seguir (SEEBER,1993,p.46), (HOFMANN-WELLENHOF, LICHTENEGGER ; COLLINS, 1994, p113):

$$d_{trop,d} = \frac{10^{-6}}{5} \frac{77,64 \frac{P}{T}}{\sin \sqrt{E^2 + 6,25}} [40136 + 148,72 (T - 273,16)]; \quad (2.41)$$

$$d_{\text{trop},w} = \frac{10^{-6} - 12,96 T + 3,718 \cdot 10^5}{5 \sin \sqrt{E^2 + 2,25}} \frac{e}{T^2} 11000; \quad (2.42)$$

$$d_{\text{trop}} = d_{\text{trop},d} + d_{\text{trop},w}, \quad (2.43)$$

onde:

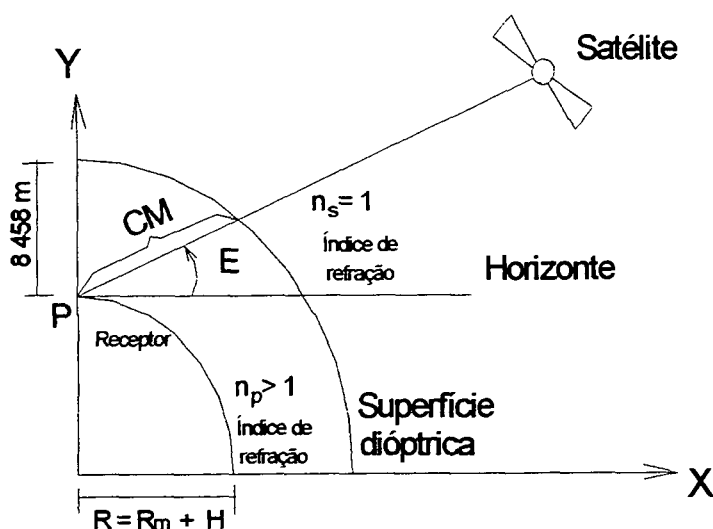
E : ângulo de elevação do satélite⁵;

T, P, e : já foram definidos com suas respectivas unidades (equação 2.39).

ANDRADE (1988, p.30-35) apresenta um modelo que substitui trechos da atmosfera por uma superfície dióptrica esférica situada a 8 458 m de altitude (fig.2.8).

O índice de refração do ponto P (n_p) é o mesmo até a superfície dióptrica e igual a 1 acima dela, $n_s=1$. O raio R_m pode ser definido para ser o raio médio de curvatura das seções normais, calculado a partir da seção normal primeiro vertical e da seção normal meridiana (ver GEMAEL, 1977, p. 3.11). H é a altitude ortométrica da antena do receptor e, E é o ângulo de elevação do satélite.

FIGURA 2.8 – ATMOSFERA SIMPLIFICADA



FONTE: ANDRADE, J. B. de. **NAVSTAR-GPS**. Curitiba, 1988. 63 f. Apostila - Curso de Pós-Graduação em Ciências Geodésicas, Universidade Federal do Paraná, p.31.

⁵ O ângulo de elevação do satélite pode ser calculado com base no sistema geodésico local ou no sistema topográfico local. Ambos os sistemas são amplamente discutidos em GARNÉS (1998a).

A refração troposférica nesse modelo é dada por:

$$d_{\text{trop}} = \text{CM} \cdot N \cdot 10^{-6}, \quad (2.44)$$

com:

N : a refratividade calculada pela equação (2.39);

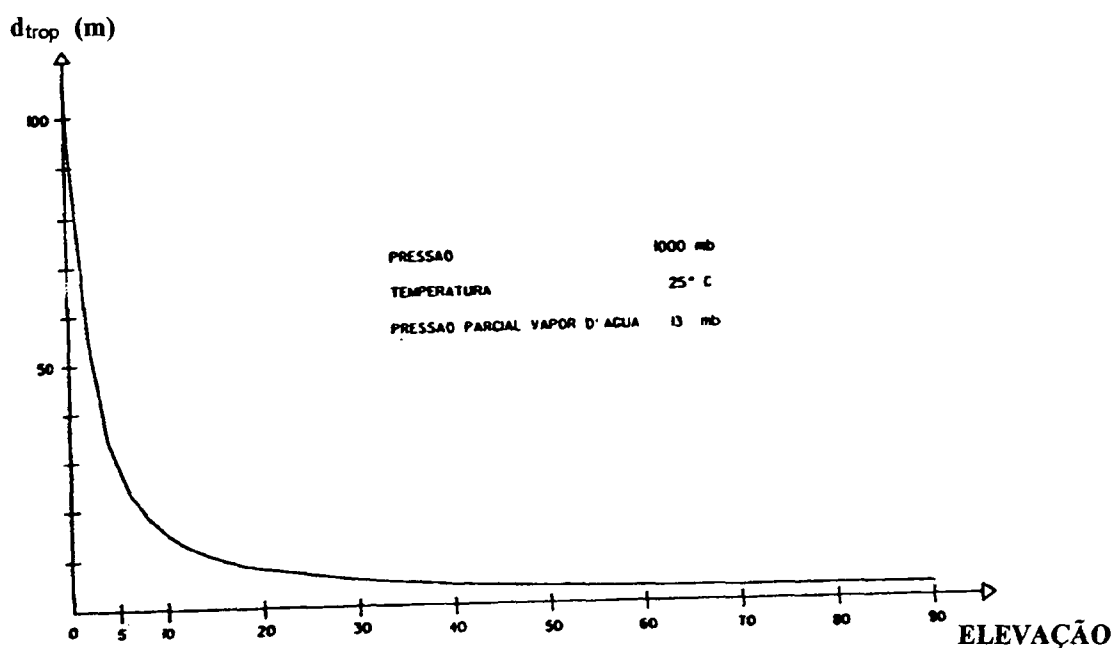
$$\text{CM} = X \cdot \sqrt{1 + \text{tg}^2 E};$$

$$X = \frac{-2R \text{tg} E + \sqrt{(2R \text{tg} E)^2 + 4(1 + \text{tg}^2 E) 8458 (2R + 8458)^2}}{2(1 + \text{tg}^2 E)};$$

$$R = R_m + H.$$

A fig.2.9 mostra o comportamento da refração troposférica para os valores da pressão $P = 1.000 \text{ mb}$, temperatura $T = 25^\circ \text{C}$ e pressão parcial do vapor d'água $e = 13 \text{ mb}$ em função do ângulo de elevação (E) do satélite.

FIGURA 2.9 – REFRAÇÃO TROPOSFÉRICA



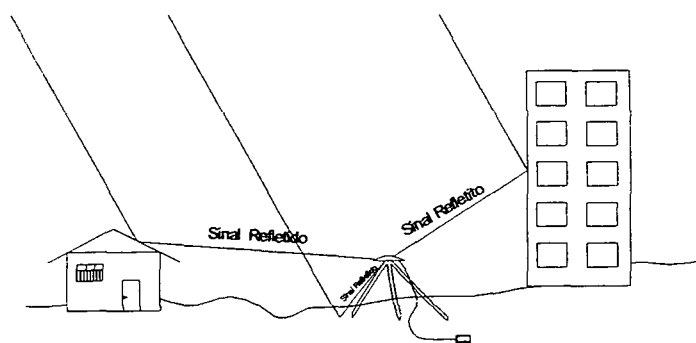
FONTE: ANDRADE, J. B. de. **NAVSTAR-GPS**. Curitiba, 1988. 63 f. Apostila - Curso de Pós-Graduação em Ciências Geodésicas, Universidade Federal do Paraná, p.36.

2.5.11 Multicaminho

É o efeito correspondente à chegada do sinal na antena do receptor, por diferentes caminhos (dois ou mais), devido à reflexão no meio (fig. 2.10).

O multicaminho na medida de distâncias usando os códigos, segundo TRANQUÍLA e CARR, (1990/91)⁶, citado por HOFMANN-WELLENHOF, LICHTENEGGER e COLLINS, (1994, p.125) podem ser agrupados nas seguintes classes: i) espalhamento de avanço difuso (*diffuse forward scattering*) em área amplamente distribuída (o sinal passa através de objetos que refletem desordenadamente), causa um erro em torno de 10 m; ii) reflexão especular a partir de objetos bem definidos nas vizinhanças da antena, causa um erro de 2 à 6 m; e iii) flutuações de baixa frequência, geralmente associado à reflexão da superfície da água, causa um erro em torno de 10 m.

FIGURA 2.10 - MULTICAMINHO



Na medida da fase o multicaminho é bem menor e atinge no máximo a quarta parte do comprimento de onda da observável. A análise desse resultado, apresentada a seguir, é clássica na literatura e pode ser encontrada em SEEBER(1993, p.307),

⁶ TRANQUILLA, J. M.; CARR, J. P. GPS multipath field observations at land and water sites. *Navigation*, v. 37, n. 4, p. 393-414, 1990/91.

LEICK (1995,p.313), HOFMANN-WELLENHOF, LICHTENEGGER e COLLINS, (1994, p.126) e LANGLEY (1998, p.141).

Considere as equações (2.45) e (2.46) como sendo respectivamente o sinal recebido diretamente na antena e o sinal refletido antes de chegar na antena:

$$A_D = A \cos \varphi ; \quad (2.45)$$

$$A_R = \alpha A \cos(\varphi + \Delta\varphi), \quad (2.46)$$

onde:

A_D : é a amplitude do sinal direto;

A_R : é a amplitude do sinal refletido;

$\alpha \in [0,1]$: fator de atenuação, com $\alpha=0$ não há reflexão e $\alpha=1$ o sinal é refletido com a mesma intensidade do sinal direto;

φ : é a fase do sinal direto; e

$\Delta\varphi$: é a mudança na fase devido ao multicaminho.

A superposição de ambos os sinais é:

$$\begin{aligned} A_\Sigma &= A_D + A_R = A \cos\varphi + \alpha A \cos(\varphi + \Delta\varphi) \\ &= (1 + \alpha \cos\Delta\varphi)A \cos\varphi - (\alpha \sin\Delta\varphi)A \sin\varphi \end{aligned} \quad (2.47)$$

O sinal (2.47) também pode ser colocado na forma:

$$\begin{aligned} A_\Sigma &= \beta A \cos(\varphi + \Theta) \\ &= (\beta \cos \Theta)A \cos\varphi - (\beta \sin \Theta) A \sin\varphi \end{aligned} \quad (2.48)$$

onde

β : atenuação do sinal superposto;

Θ : fase do sinal superposto por causa do multicaminho.

Comparando as equações (2.48) e (2.47) tira-se:

$$\beta \sin \Theta = \alpha \sin \Delta\varphi \quad (2.49)$$

$$\beta \cos \Theta = 1 + \alpha \cos \Delta\varphi \quad (2.50)$$

Dividindo a equação (2.49) pela equação (2.50) tem-se:

$$\operatorname{tg} \Theta = \frac{\alpha \sin \Delta\varphi}{1 + \alpha \cos \Delta\varphi} \quad (2.51)$$

Quadrando as equações (2.49) e (2.50) e somando resulta

$$\beta = \sqrt{1 + \alpha^2 + 2\alpha \cos \Delta\varphi} \quad (2.52)$$

A máxima intensidade que o sinal refletido pode ter é quando $\alpha=1$. Introduzindo esse valor nas equações (2.51) e (2.52) tem-se, respectivamente:

$$\operatorname{tg} \Theta = \frac{\sin \Delta\varphi}{1 + \cos \Delta\varphi} = \operatorname{tg} \frac{\Delta\varphi}{2} \quad \Rightarrow \quad \Theta = \frac{\Delta\varphi}{2} \quad (2.53)$$

$$\beta = \sqrt{2(1 + \cos \Delta\varphi)} = 2 \cos \frac{\Delta\varphi}{2} \quad (2.54)$$

Analisando os valores para as fases $\Delta\varphi$ e Θ , o máximo valor para Θ é alcançado em $\Theta = 90^\circ$ o que corresponde a $\lambda/4$. Para a portadora L_1 , $\lambda_1=19,05$ cm, implica em 4,8 cm e para a portadora L_2 , $\lambda_2=24,45$ cm, implica em 6,1 cm.

O multicaminho pode ser determinado usando medidas das duas portadoras L_1 e L_2 . A idéia é que a refração troposférica, erros dos relógios e efeitos relativísticos influenciam o código e a fase pelas mesmas quantidades nas duas portadoras, ao passo que a refração ionosférica e o multicaminho dependem da frequência. Assim, se for formada a combinação linear livre da ionosfera (*ionospheric-free*) para o código e para a fase da portadora e, realizada a diferença, a parte sistemática que sobra é devido ao multicaminho (HOFMANN-WELLENHOF, LICHTENEGGER e COLLINS, 1994, p.124).

O multicaminho além de influenciar na acurácia do posicionamento pode fazer com que haja perdas de ciclo (*cycle slip*) na fase da portadora. Segundo LEICK (1995, p.314), vários fatores fazem com que o efeito do multicaminho seja um erro difícil de se tratar. No entanto, alguns cuidados podem ser tomados a fim de se minimizar esse efeito, são eles (SEEBER, 1993,p.309), (LANGLEY, 1998, p.142):

- selecionar cuidadosamente os lugares para colocar a antena (evitar lugares onde o sinal possa ter reflexão);
- usar antenas com características apropriadas para tratar do multicaminho;
- uso de material absorvente próximo das antenas;
- uso de receptores com discriminante de multicaminho.

2.6 EQUAÇÕES DE OBSERVAÇÃO

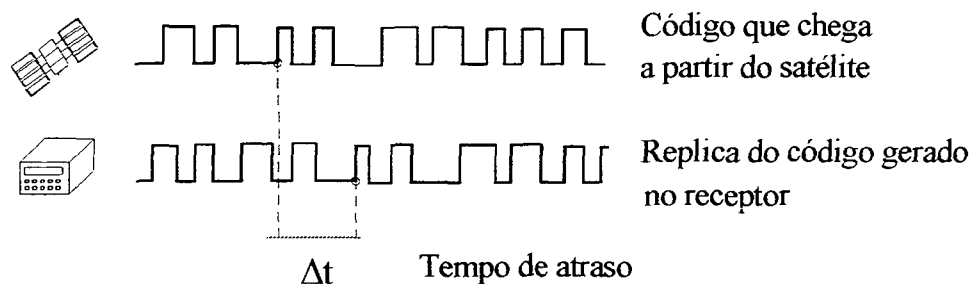
2.6.1 Equação de Observação do Código

Nos receptores são geradas réplicas idênticas aos códigos transmitidos pelos satélites. Um detector de correlação identifica quando o código gerado pelo receptor se alinha (máxima correlação) com o código que chega transmitido pelo satélite (fig. 2.11). O intervalo de tempo necessário para a máxima correlação é então registrado na escala de tempo do relógio do receptor. O código transmitido pelo satélite é gerado com base no relógio do satélite, o código gerado pelo receptor é produzido com base no relógio do receptor, e, ambos os relógios, em geral, não são sincronizados. Se o sincronismo entre os relógios do satélite e receptor acontecesse, o tempo de deslocamento do código para obtenção da máxima correlação, corresponderia ao tempo em que a onda portadora levou para percorrer a distância entre as antenas do satélite e do receptor (a menos de atrasos dos *hardwares*).

Chamando t_R , o instante de recepção do sinal do satélite no relógio do receptor e, t^S , o instante de transmissão do sinal no relógio do satélite, o intervalo de tempo medido é:

$$\Delta t = t_R - t^S. \quad (2.55)$$

FIGURA 2.11 – MEDIDA COM O CÓDIGO



FONTE: WELLS, D.; BECK, N.; DELIKARAOGLOU, D.; KLEUSBERG, A.; KRAKIWSKY, E. J.; LACHAPELLE, G.; LANGLEY, R. B.; NAKIBOGLU, M. SCHWARZ, K. P.; TRANQUILLA, J. M.; VANÍČEK, P. **Guide to GPS positioning**. Canadá: Canadian GPS Associates, Fredericton, N. B., 1986.1986, p. 4.14.

O produto de Δt pela velocidade da luz no vácuo ($c = 299\,792\,458$ m/s), produz uma distância (P_R) entre receptor e satélite afetada de todo tipo de erro inerente ao processo (ver seção 2.5). Por causa destes erros essa distância recebeu a denominação em inglês de *pseudo-range*. No Brasil, o termo *pseudo-range* foi traduzido por pseudodistância, e vem sendo utilizado com essa denominação desde 1988 com os trabalhos de ANDRADE (1988), SANTOS (1990), TIERRA CRIOLLO (1993), KRUEGER (1994) e seguindo por outros autores. A equação é:

$$P_R = c \Delta t. \quad (2.56)$$

A pseudodistância (P_R), é a medida que o receptor registra quando utiliza os códigos P(Y) ou C/A. Para relacionar a pseudodistância medida com as coordenadas da estação a determinar, é necessário introduzir as seguintes relações (GURTNER;

MADER; MACARTHUR, 1989), (GURTNER; MADER, 1990):

$$t_R = t_R^{\text{GPS}} + dt_R; \quad (2.57)$$

$$t^S = t_{\text{GPS}}^S + dt^S, \quad (2.58)$$

onde:

t_R^{GPS} : instante de recepção do sinal no sistema de tempo GPS;

t_{GPS}^S : instante de transmissão do sinal no sistema de tempo GPS;

dt_R : estado do relógio do receptor em relação ao tempo GPS;

dt^S : estado do relógio do satélite em relação ao tempo GPS.

Substituindo as equações (2.57) e (2.58) na (2.55) e considerando a (2.56), tem-se:

$$P_R = c (t_R^{\text{GPS}} - t_{\text{GPS}}^S) + c (dt_R - dt^S). \quad (2.59)$$

A diferença $(t_R^{\text{GPS}} - t_{\text{GPS}}^S) = \tau$, de acordo com TEUNISSEN e KLEUSBERG (1998, p. 190) pode ser explicitada por:

$$\tau = \frac{1}{c} d_R^{\text{hard}} + \delta\tau + \frac{1}{c} d_{\text{hard}}^S, \quad (2.60)$$

onde:

d_R^{hard} : atraso do hardware do receptor;

$\delta\tau$: tempo de propagação do sinal entre a antena transmissora e antena receptora;

d_{hard}^S : atraso do hardware do satélite.

O intervalo $\delta\tau$, é dado por:

$$\delta\tau = \frac{1}{c} (\rho_R^S + d_{\text{mult}} + d_{\text{ion}} + d_{\text{trop}}), \quad (2.61)$$

onde:

ρ_R^S : distância entre as antenas do satélite e do receptor no instante t_R ;

d_{mult} : multicaminho; e

d_{ion} , d_{trop} : refração ionosférica e troposférica, respectivamente.

Levando-se a (2.61) em (2.60), e posteriormente em (2.59) com o decréscimo do resíduo ε , tem-se:

$$P_R = \rho_R^S + c dt_R - c dt^S + d_{trop} + d_{ion} + d_{mult} + d_R^{hard} + d_{hard}^S - \varepsilon . \quad (2.62)$$

Nesse trabalho, no entanto, é utilizada a equação da pseudodistância numa forma simplificada, uma vez que não se pretende determinar d_{mult} , d_R^{hard} , d_{hard}^S

$$P_R = \rho_R^S + c dt_R - c dt^S + d_{trop} + d_{ion} - \varepsilon . \quad (2.63)$$

A distância ρ_R^S pode ser expressa em função das coordenadas da antena do satélite e receptor, por:

$$\rho_R^S = \sqrt{(x^S - x_R)^2 + (y^S - y_R)^2 + (z^S - z_R)^2} , \quad (2.64)$$

com:

x^S, y^S, z^S : coordenadas geocêntricas da antena do satélite; e

x_R, y_R, z_R : coordenadas geocêntricas da antena do receptor.

2.6.2 Equação de Observação da Fase da Portadora

A observação da fase da portadora é a diferença entre a fase da portadora do sinal gerado pelo oscilador interno do receptor e a fase da portadora do sinal que chega transmitido a partir do satélite.

A observação é registrada na escala de tempo de relógio do receptor e

representada pela equação:

$$\varphi_R^s(t_R) = \varphi_R(t_R) - \varphi^s(t_R), \quad (2.65)$$

onde:

t_R : instante de recepção do sinal no relógio do receptor;

$\varphi_R^s(t_R)$: fase da portadora medida no instante t_R ;

$\varphi_R(t_R)$: fase da portadora do sinal do receptor no instante t_R ;

$\varphi^s(t_R)$: fase da portadora do sinal recebido do satélite no instante t_R .

A fase $\varphi_R^s(t_R)$ é a medida da diferença $[\varphi_R(t_R) - \varphi^s(t_R)]$ em até um ciclo e, conforme mencionado na seção 1.1, se o sinal do satélite for continuamente rastreado a partir de uma primeira medida t_0 , o receptor mantém a contagem de ciclos inteiros até a época t_R e, com isso, pode-se determinar a distância que o satélite se aproximou ou se afastou do receptor desde a época t_0 , na direção satélite/receptor (ver fig. 1.1).

O desenvolvimento da equação (2.65), é baseado na equivalência da fase recebida ($\varphi^s(t_R)$), com a fase transmitida ($\varphi_T^s(t_R - \tau)$) pelo satélite τ segundos antes (LEICK, 1995, p. 257). A equação é:

$$\varphi^s(t_R) = \varphi_T^s(t_R - \tau), \quad (2.66)$$

ou ainda:

$$\varphi^s(t_R) \equiv \varphi_T^s(t_R - \tau) = \varphi_{TT}^s(t^s), \quad (2.67)$$

onde: τ é o tempo de propagação do sinal.

A equação (2.67) merece melhores esclarecimentos, ou seja, a fase recebida $\varphi^s(t_R)$ no instante t_R do relógio do receptor, é equivalente a fase transmitida $\varphi_T^s(t_R - \tau)$ pelo satélite no instante $(t_R - \tau)$ do relógio do receptor e, corresponde à fase do satélite $\varphi_{TT}^s(t^s)$ no instante t^s de transmissão do relógio do satélite.

Usando as relações definidas nas equações (2.57) e (2.58) e considerando a estabilidade dos osciladores do receptor e do satélite, pode-se escrever:

$$\varphi_R(t_R) = \varphi_R(t_R^{\text{GPS}} + dt_R) = \varphi_R(t_R^{\text{GPS}}) + f_R dt_R, \quad (2.68)$$

$$\varphi_{\text{TT}}^s(t^S) = \varphi_{\text{TT}}^s(t_{\text{GPS}}^S + dt^S) = \varphi_{\text{TT}}^s(t_{\text{GPS}}^S) + f_S dt^S, \quad (2.69)$$

onde f_S e f_R são as frequências dos sinais do satélite e do receptor, respectivamente.

O instante de transmissão do sinal no sistema de tempo GPS é igual ao instante de recepção do sinal no sistema de tempo GPS, menos o tempo de propagação do sinal, $t_{\text{GPS}}^S = t_R^{\text{GPS}} - \tau$. Assim:

$$\varphi_{\text{TT}}^s(t_{\text{GPS}}^S) = \varphi_{\text{TT}}^s(t_R^{\text{GPS}}) - f_S \tau. \quad (2.70)$$

Substituindo essa última equação em (2.69), vem:

$$\varphi_{\text{TT}}^s(t^S) = \varphi_{\text{TT}}^s(t_R^{\text{GPS}}) - f_S \tau + f_S dt^S. \quad (2.71)$$

Tomando em conta a equação (2.67), e levando as equações (2.68) e (2.71) na equação (2.65), resulta:

$$\varphi_R^s(t_R) = \varphi_R(t_R^{\text{GPS}}) - \varphi_{\text{TT}}^s(t_R^{\text{GPS}}) + f_S \tau + f_R dt_R - f_S dt^S. \quad (2.72)$$

A diferença $(\varphi_R(t_R^{\text{GPS}}) - \varphi_{\text{TT}}^s(t_R^{\text{GPS}}))$, quando já considerados os termos $f_R dt_R$ e $f_S dt^S$, pode ser representada por (HOFMANN-WELLENHOF; LICHTENEGGER; COLLINS, 1994, p.90):

$$\varphi_R(t_R^{\text{GPS}}) - \varphi_{\text{TT}}^s(t_R^{\text{GPS}}) = (f_R - f_S) t_R^{\text{GPS}}, \quad (2.73)$$

e negligenciada, porque a diferença está na ordem de 10^{-4} ciclos, ficando abaixo do

nível de ruído da portadora.

O tempo de propagação do sinal τ para a fase, é explicitado de modo semelhante ao do código, equações (2.60) e (2.61). A observação é que a refração ionosférica deve ser precedida pelo sinal negativo, porque a fase da portadora do sinal transmitido adianta quando o sinal atravessa a ionosfera, ao passo que para o código, atrasa (ver seção 2.5.9).

Reescrevendo as equações (2.60) e (2.61) para a fase, tem-se:

$$\tau = \frac{1}{c} d_R^{\text{hard}} + \delta\tau + \frac{1}{c} d_{\text{hard}}^s, \quad (2.74)$$

com

$$\delta\tau = \frac{1}{c} (\rho_R^s + d_{\text{mult}} - d_{\text{ion}} + d_{\text{trop}}). \quad (2.75)$$

Levando (2.75) em (2.74) e com as considerações sobre a expressão (2.73), isto é, $f_R = f_S = f$, obtém-se para (2.72), a expressão:

$$\varphi_R^s(t_R) = \frac{f}{c} \rho_R^s + f dt_R - f dt^s + \frac{f}{c} (d_{\text{trop}} - d_{\text{ion}} + d_{\text{mult}} + d_R^{\text{hard}} + d_{\text{hard}}^s). \quad (2.76)$$

O lado direito representa o total de ciclos entre receptor e satélite na época t_R e, o lado esquerdo, o número de ciclos contados a partir da época t_0 ($N_R^s(t_R) + \varphi_R^s(t_R)$), conforme é ilustrado na fig.1.1. Para retirar a “inconsistência” da equação (2.76), é necessário subtrair do lado direito o número inteiro de ciclos ($N_R^s(t_0)$), resultando a equação:

$$\varphi_R^s(t_R) = \frac{f}{c} \rho_R^s + f dt_R - f dt^s + \frac{f}{c} (d_{\text{trop}} - d_{\text{ion}} + d_{\text{mult}} + d_R^{\text{hard}} + d_{\text{hard}}^s) - N_R^s \quad (2.77)$$

O número inteiro de ciclos iniciais $N_R^s = N_R^s(t_0)$ é a ambigüidade, e por ser a incógnita fundamental da equação, deve ser muito bem determinada.

Lembrando que λ tem unidade m/ciclo, basta multiplicar ambos os lados da

equação (2.77) por λ , para passá-la a metros. Assim,

$$\Phi_R^s(t_R) = \rho_R^s + c dt_R - c dt^s + d_{trop} - d_{ion} + d_{mult} + d_R^{hard} + d_{hard}^s - \lambda N_R^s, \quad (2.78)$$

onde:

$$\Phi_R^s(t_R) = \lambda \varphi_R^s(t_R).$$

Autores como, por exemplo, LEICK (1995, p. 259), TEUNISSEN e KLEUSBERG, (1998, p. 194), apresentam a equação de observação da fase da portadora com mais alguns termos em consideração. Nesse trabalho, a equação será utilizada numa forma ainda mais simplificada, uma vez que os erros devido ao multicaminho (d_{mult}) e atrasos dos *hardwares* (d_R^{hard} e d_{hard}^s), serão considerados implícitos no resíduo (ε). A equação simplificada, então, torna-se:

$$\Phi_R^s = \rho_R^s + c dt_R - c dt^s + d_{trop} - d_{ion} - \lambda N_R^s - \varepsilon. \quad (2.79)$$

Na equação (2.79), foi omitido o instante (t_R) para simplificar a notação.

2.6.3 Combinação Linear das Observáveis

Com as medidas da pseudodistância e da fase da portadora, escreveu-se a equação de observação para o código (equação 2.63), e para a fase da portadora (equação 2.79). A partir da combinação linear delas, pode-se derivar uma infinidade de outras equações de observação. Algumas apresentam características bastante desejáveis para o processamento GPS.

Dentre os procedimentos de obtenção das combinações lineares, destaca-se:

- combinação linear das observáveis em L_1 e L_2 ;
- simples diferença entre receptores e satélites;

- dupla diferença; e
- tripla diferença.

2.6.3.1 Combinação linear das observáveis em L_1 e L_2

O receptor de dupla frequência pode medir a fase da portadora em um determinado instante para a $L_1(\varphi_1)$ e para a $L_2(\varphi_2)$. No caso das pseudodistâncias em L_2 é necessário um receptor com acesso ao código $P(Y)$, ou, um que obtenha a pseudodistância via alguma técnica especial (ver quadro 2.1).

No caso da fase da portadora (em ciclos), a combinação linear é escrita como:

$$\varphi_{n,m} = n \varphi_1 + m \varphi_2, \quad (2.80)$$

com $n, m \in \mathbb{R}$.

A frequência (f), comprimento de onda (λ), ambigüidade (N) e ruído (*ruído*) para a observável resultante ($\varphi_{n,m}$), são representados respectivamente por (SEEBER, 1993, p.258):

$$f_{n,m} = n f_1 + m f_2, \quad (2.81)$$

$$\lambda_{n,m} = \frac{c}{n f_1 + m f_2}, \quad (2.82)$$

$$N_{n,m} = n N_1 + m N_2, \quad (2.83)$$

$$Ruído_{(n,m)} = 1,592 \cdot 10^{-2} \lambda_{n,m} \sqrt{n^2 + m^2}, \quad (2.84)$$

onde:

$$1,592 \cdot 10^{-2} = \frac{3,03 \text{ mm}}{190,5 \text{ mm}} : \text{é o ruído da fase em } L_1 \text{ (ciclos)}.$$

A tab 2.6 contém algumas das possíveis combinações lineares formadas com as fases em L_1 e L_2 . As que estão selecionadas nas áreas mais escuras, são as mais populares. L_w é a *wide lane*, L_N é a *narrow lane*, L_0 é a combinação livre da ionosfera (*ionospheric free*) e L_I é a combinação com a ionosfera completa.

TABELA 2.6 – COMBINAÇÕES LINEARES DA FASE DAS PORTADORAS

IDENTIFICAÇÃO	n	m	$\lambda_{n,m}$ (cm)	Ruído _(n,m) (mm)
L_1	1	0	19,0	3,0
L_2	0	1	24,4	3,9
L_w	1	-1	86,2	19,4
L_N	1	1	10,7	2,1
$L_{1,2}$	-1	2	34,1	12,1
$L_{2,-1}$	2	-1	15,6	5,5
$L_{3,-2}$	3	-2	13,2	7,6
$L_{4,-3}$	4	-3	11,4	9,1
$L_{9,-7}$	9	-7	5,4	9,7
$L_{5,-4}$	5	-4	10,1	10,3
$L_{6,-5}$	6	-5	9,0	11,2
$L_{7,-6}$	7	-6	8,2	12,0
$L_{8,-7}$	8	-7	7,5	12,6
$L_{9,-8}$	9	-8	6,9	13,2
$L_{10,-9}$	10	-9	6,4	13,6
L_0	-	-	≈ 5,4	10,0
L_I	-	-	≈ 10,7	20,0

FONTE: Adaptado de SEEGER, G. **Satellite geodesy: foundations, methods, and applications**. Berlin: Walter de Gruyter, 1993, p.259.

As combinações L_0 e L_I , cujos valores de n e m não foram indicadas na tabela, são:

$$L_0 = \frac{L_w + L_N}{2}, \quad (2.85)$$

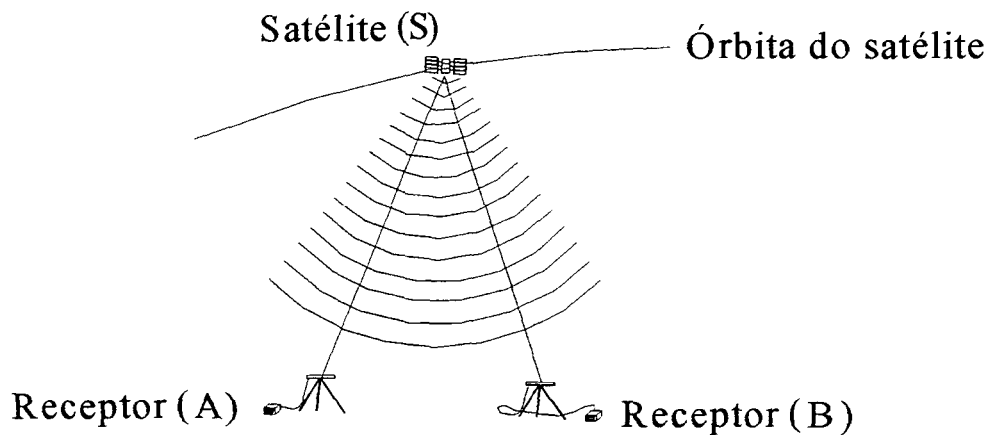
$$L_I = L_N - L_w. \quad (2.86)$$

2.6.3.2 Equação de observação da simples diferença

A simples diferença, é a combinação linear que envolve as equações de observação de dois receptores (em estações distintas A e B) a um mesmo satélite (S), ou então, envolve as equações de observação de dois satélites (S e P) a um único receptor (A).

A combinação envolvendo dois receptores e um satélite é chamada de simples diferença entre receptores e, a combinação envolvendo dois satélites e um receptor é chamada de simples diferença entre satélites (WELLS et al.,1986,p.8.6-8.7). A fig. (2.12) ilustra a geometria da simples diferença entre receptores.

FIGURA 2.12 – GEOMETRIA DA SIMPLES DIFERENÇA ENTRE RECEPTORES



Denotando-se a equação de observação da pseudodistância (2.63), para os receptores A e B e satélite S, como:

$$P_A^S = \rho_A^S + c dt_A - c dt^S + d_{trop(A)}^S + d_{ion(A)}^S - \varepsilon_A^S ; \quad (2.87)$$

$$P_B^S = \rho_B^S + c dt_B - c dt^S + d_{trop(B)}^S + d_{ion(B)}^S - \varepsilon_B^S , \quad (2.88)$$

e, subtraindo (2.87) de (2.88), obtém-se a equação de observação da simples diferença de pseudodistâncias:

$$P_{AB}^s = P_B^s - P_A^s = \rho_B^s - \rho_A^s + c(dt_B - dt_A) + (d_{trop(B)}^s - d_{trop(A)}^s) + (d_{ion(B)}^s - d_{ion(A)}^s) - (\varepsilon_B^s - \varepsilon_A^s)$$

$$P_{AB}^s = \rho_{AB}^s + cdt_{AB} + d_{trop(AB)}^s + d_{ion(AB)}^s - \varepsilon_{AB}^s. \quad (2.89)$$

Se o estado do relógio do satélite (dt^s) fosse constante, a simples diferença entre receptores cancelaria esse erro. No entanto, conforme mostrado na equação (2.12), dt^s não é constante, exigindo assim, a simultaneidade entre as observações. Contudo, mesmo sendo simultâneas as observações nos receptores A e B, por causa da diferença de distância entre elas e os satélites, os sinais recebidos simultaneamente nesses receptores, teriam sido transmitidos em instantes diferentes, e, com o aumento da linha de base, torna-se necessário fazer as correções apropriadas.

Conforme mencionado na seção (2.5.9), a refração ionosférica para a simples diferença ($d_{ion(AB)}^s$), tem ordem de grandeza de 1ppm em linhas de base curta (costuma-se dizer que a linha de base é curta, quando seu comprimento não ultrapassa a 10 km). Deste modo, dependendo da aplicação, pode ser desprezada.

As refrações troposféricas ($d_{trop(B)}^s$ e $d_{trop(A)}^s$) podem ser corrigidas de acordo com alguns dos modelos apresentados anteriormente (seção 2.5.10). Para linhas de base curta, onde as condições atmosféricas de temperatura e pressão são semelhantes, a diferença $d_{trop(AB)}^s$ pode ser negligenciada.

A equação (2.89), se refere à equação de observação da simples diferença das pseudodistâncias medidas com o código. Procedendo-se de forma análoga com a equação (2.79), obtém-se a equação de observação da simples diferença de fase

$$\Phi_{AB}^s = \rho_{AB}^s + cdt_{AB} + d_{trop(AB)}^s - d_{ion(AB)}^s - \lambda N_{AB}^s - \varepsilon_{AB}^s \quad (2.90)$$

onde:

$$\Phi_{AB}^s = \Phi_B^s - \Phi_A^s;$$

$$\rho_{AB}^s = \rho_B^s - \rho_A^s;$$

$$dt_{AB} = dt_B - dt_A;$$

$$d_{trop(AB)}^S = d_{trop(B)}^S - d_{trop(A)}^S ;$$

$$d_{ion(AB)}^S = d_{ion(B)}^S - d_{ion(A)}^S ;$$

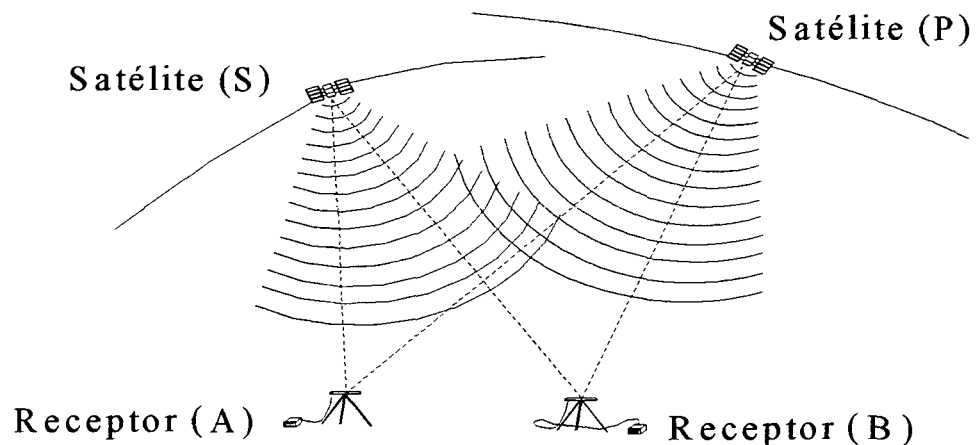
$$\varepsilon_{AB}^S = \varepsilon_B^S - \varepsilon_A^S .$$

A simples diferença entre satélites é obtida da mesma forma da simples diferença entre receptores. A diferença entre elas é que o erro a ser cancelado, é o erro do relógio do receptor (dt_R). Os demais erros não são cancelados porque o caminho de propagação do sinal de dois satélites, até um mesmo receptor, é bastante diferente.

2.6.3.3 Equação de observação da dupla diferença

A combinação da dupla diferença envolve dois receptores (A e B) e dois satélites (S e P), conforme ilustra a fig. 2.13.

FIGURA 2.13 – GEOMETRIA DA DUPLA DIFERENÇA



Para o caso das pseudodistâncias, a dupla diferença é obtida subtraindo duas equações da simples diferença (2.89), uma para o satélite S e outra para o satélite P. Sendo simultâneas as observações, a simples diferença dos erros dos relógios dos receptores (dt_{AB}) é cancelada.

Considerando a simples diferença de pseudodistância para o satélite P:

$$P_{AB}^P = \rho_{AB}^P + cdt_{AB} + d_{trop(AB)}^P + d_{ion(AB)}^P - \varepsilon_{AB}^P, \quad (2.91)$$

e subtraindo a equação (2.89) da equação (2.91), tem-se:

$$P_{AB}^{SP} = \rho_{AB}^{SP} + d_{trop(AB)}^{SP} + d_{ion(AB)}^{SP} - \varepsilon_{AB}^{SP}, \quad (2.92)$$

onde:

$$\begin{aligned} P_{AB}^{SP} &= P_B^P - P_A^P - P_B^S + P_A^S; \\ \rho_{AB}^{SP} &= \rho_B^P - \rho_A^P - \rho_B^S + \rho_A^S; \\ d_{trop(AB)}^{SP} &= d_{trop(B)}^P - d_{trop(A)}^P - d_{trop(B)}^S + d_{trop(A)}^S; \\ d_{ion(AB)}^{SP} &= d_{ion(B)}^P - d_{ion(A)}^P - d_{ion(B)}^S + d_{ion(A)}^S; \\ \varepsilon_{AB}^{SP} &= \varepsilon_B^P - \varepsilon_A^P - \varepsilon_B^S + \varepsilon_A^S. \end{aligned}$$

Procedendo de forma análoga com as equações da simples diferença de fase, equação (2.90), obtém-se a equação da dupla diferença de fase:

$$\Phi_{AB}^{SP} = \rho_{AB}^{SP} + d_{trop(AB)}^{SP} - d_{ion(AB)}^{SP} - \lambda N_{AB}^{SP} - \varepsilon_{AB}^{SP}, \quad (2.93)$$

com:

$$\begin{aligned} \Phi_{AB}^{SP} &= \Phi_B^P - \Phi_A^P - \Phi_B^S + \Phi_A^S; \\ N_{AB}^{SP} &= N_B^P - N_A^P - N_B^S + N_A^S; \\ \rho_{AB}^{SP} &= \rho_B^P - \rho_A^P - \rho_B^S + \rho_A^S; \\ d_{trop(AB)}^{SP} &= d_{trop(B)}^P - d_{trop(A)}^P - d_{trop(B)}^S + d_{trop(A)}^S; \\ d_{ion(AB)}^{SP} &= d_{ion(B)}^P - d_{ion(A)}^P - d_{ion(B)}^S + d_{ion(A)}^S; \\ \varepsilon_{AB}^{SP} &= \varepsilon_B^P - \varepsilon_A^P - \varepsilon_B^S + \varepsilon_A^S. \end{aligned}$$

2.6.3.4 Equação de observação da tripla diferença

A tripla diferença é obtida realizando a diferença entre duas equações da dupla diferença em épocas distintas. É usada principalmente por eliminar a dupla

diferença de ambigüidades da fase da portadora (se não ocorrer perdas de ciclo entre épocas, as ambigüidades não mudam).

Considerando duas épocas t_1 e t_2 , e escrevendo a dupla diferença de fase para cada uma delas, tem-se:

$$\Phi_{AB}^{SP}(t_1) = \rho_{AB}^{SP}(t_1) + d_{trop(AB)}^{SP}(t_1) - d_{ion(AB)}^{SP}(t_1) - \lambda N_{AB}^{SP}(t_1) - \varepsilon_{AB}^{SP}(t_1); \quad (2.94)$$

$$\Phi_{AB}^{SP}(t_2) = \rho_{AB}^{SP}(t_2) + d_{trop(AB)}^{SP}(t_2) - d_{ion(AB)}^{SP}(t_2) - \lambda N_{AB}^{SP}(t_2) - \varepsilon_{AB}^{SP}(t_2). \quad (2.95)$$

Subtraindo a equação (2.95) da equação (2.94), resulta na equação da tripla diferença de fase:

$$\begin{aligned} \Phi_{AB}^{SP}(t_2 - t_1) = & [\rho_{AB}^{SP}(t_1) - \rho_{AB}^{SP}(t_2)] + [d_{trop(AB)}^{SP}(t_1) - d_{trop(AB)}^{SP}(t_2)] - [d_{ion(AB)}^{SP}(t_1) - \\ & d_{ion(AB)}^{SP}(t_2)] - \lambda [N_{AB}^{SP}(t_1) - N_{AB}^{SP}(t_2)] - [\varepsilon_{AB}^{SP}(t_1) - \varepsilon_{AB}^{SP}(t_2)]. \end{aligned} \quad (2.96)$$

Trocando o sinal que precede o termo da refração ionosférica, obtém-se a equação da tripla diferença de pseudodistâncias:

$$\begin{aligned} P_{AB}^{SP}(t_2 - t_1) = & [\rho_{AB}^{SP}(t_1) - \rho_{AB}^{SP}(t_2)] + [d_{trop(AB)}^{SP}(t_1) - d_{trop(AB)}^{SP}(t_2)] + [d_{ion(AB)}^{SP}(t_1) - \\ & d_{ion(AB)}^{SP}(t_2)] - \lambda [N_{AB}^{SP}(t_1) - N_{AB}^{SP}(t_2)] - [\varepsilon_{AB}^{SP}(t_1) - \varepsilon_{AB}^{SP}(t_2)]. \end{aligned} \quad (2.97)$$

3. AJUSTAMENTO DE OBSERVAÇÕES

As observações (ou medidas) quando realizadas sobre alguma grandeza física ou geométrica, têm o propósito de quantificação. Se essa quantificação carece de unicidade, seja pela propriedade estocástica das observações, seja pela falta de rigidez do modelo adotado, deve-se aplicar o conceito de ajustamento.

Em Geodésia, os modelos funcionais para ajustar as observações, em geral, têm uma das 3 formas:

$$\mathbf{L}_a = \mathbf{F}(\mathbf{x}_a) , \quad (3.1)$$

$$\mathbf{G}(\mathbf{L}_a) = \mathbf{0} , \quad (3.2)$$

$$\mathbf{H}(\mathbf{L}_a, \mathbf{x}_a) = \mathbf{0} , \quad (3.3)$$

onde:

$\mathbf{L}_a \in \mathbf{R}^m$: vetor de observações ajustadas;

$\mathbf{x}_a \in \mathbf{R}^n$: vetor de parâmetros ajustados.

No capítulo 2 (ver seção 2.6) foram desenvolvidas as equações de observação GPS. Essas equações correspondem ao modelo (3.1), de maneira que será o único a ser abordado nesse trabalho. Tratamento dos modelos (3.2) e (3.3) podem ser encontrados, por exemplo, em MIKHAIL e ACKERMAN (1976), DALMOLIN (1976), MIKHAIL e GRACIE (1981), GEMAEL (1994) e LEICK (1995).

O ajustamento de observações pode ser dividido em duas fases: a primeira é a solução de um problema matemático de otimização resultante da imposição de algum critério; e, a segunda é a validação estatística do modelo adotado, bem como dos resultados encontrados, mediante testes de hipóteses sobre a distribuição de probabilidades de alguns parâmetros estimados.

O critério adotado para a formulação do problema de otimização, será o de

mínimos quadrados, cuja concepção inicial foi desenvolvida independentemente pelos matemáticos Karl Friederich Gauss em 1795 e Adrien Marie Legendre em 1805 (MIKAIL; ACHERMAN, 1976, p. 48).

Na seção 3.1 e respectivas subseções, serão apresentados os conceitos sobre os problemas de mínimos quadrados, assim como, os métodos de solução. Na seção 3.2 e respectivas subseções, serão apresentados os conceitos estatísticos envolvidos no problema.

3.1 PROBLEMA DE MÍNIMOS QUADRADOS

No modelo (3.1), o vetor de observações ajustadas (\mathbf{L}_a) é explicitamente função dos parâmetros ajustados (\mathbf{x}_a). A função de transformação $\mathbf{F}: \mathbf{R}^n \rightarrow \mathbf{R}^m$ poderá ser linear ou não-linear, e se for não-linear, poderá ser linearizada em algum ponto \mathbf{x}_0 . Por essa razão, e pelas interpretações geométricas que o problema oferece, é conveniente iniciar considerando $\mathbf{F}: \mathbf{R}^n \rightarrow \mathbf{R}^m$ uma transformação linear.

3.1.1 Problema de Mínimos Quadrados Linear

Reescrevendo-se o modelo (3.1) como um sistema de equações lineares:

$$\mathbf{L} = \mathbf{Ax}, \quad (3.4)$$

onde:

$\mathbf{L} \in \mathbf{R}^m$: vetor de observações;

$\mathbf{A} \in \mathbf{R}^{m \times n}$: matriz de coeficientes;

$\mathbf{x} \in \mathbf{R}^n$: vetor dos parâmetros a estimar.

Supondo que $\text{posto}(\mathbf{A}) = r$, a tab.3.1 relaciona os quatro subespaços fundamentais da matriz \mathbf{A} , com suas respectivas simbologias e dimensões.

TABELA 3.1 – SUBESPAÇOS FUNDAMENTAIS DA MATRIZ A

SÍMBOLO	SIGNIFICADO	DIMENSÃO	COMPLEMENTAR
$\Re(A)$	Espaço coluna de A	r	$\Re(A)$ e $\aleph(A^T)$ $\Re(A)^\perp = \aleph(A^T)$
$\aleph(A^T)$	Espaço nulo de A^T	$m-r$	
$\Re(A^T)$	Espaço linha de A	r	$\Re(A^T)$ e $\aleph(A)$ $\Re(A^T)^\perp = \aleph(A)$
$\aleph(A)$	Espaço nulo de A	$n-r$	

NOTA: Para maiores detalhes sobre os subespaços consultar STRANG, G. **Linear algebra and its applications**. Orlando: Academic Press, (1980, p. 75, p.87)

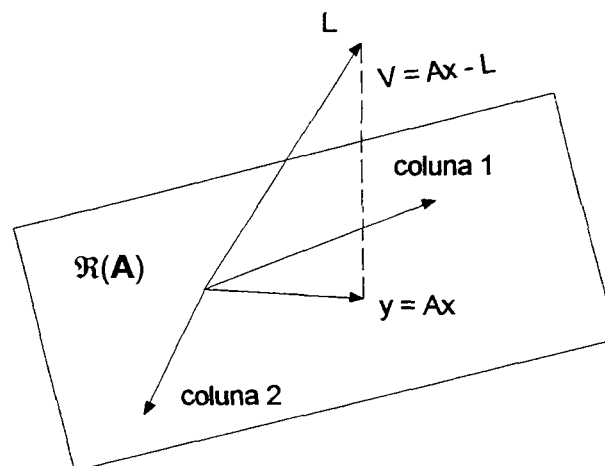
O vetor $L \in R^m$ é formado por observações e assim sujeito às flutuações probabilísticas inerentes as mesmas, para qualquer $x \in R^n$ o vetor $y = Ax \in \Re(A)$, e sendo $m > n$, $\Re(A)$ é subespaço próprio do R^m . Em geral $y \neq L$ e o sistema (3.4) é inconsistente.

Uma solução aproximada para o sistema inconsistente (3.4), pode ser encontrada resolvendo-se o problema de mínimos quadrados linear:

$$\min_{x \in R^n} \frac{1}{2} \|Ax - L\|_2^2. \quad (3.5)$$

A fig.3.1 ilustra a geometria do problema de mínimos quadrados linear, com $A \in R^{3 \times 2}$.

FIGURA 3.1 – INTERPRETAÇÃO GEOMÉTRICA DO PROBLEMA DE MÍNIMOS QUADRADOS LINEAR



A solução, no sentido de mínimos quadrados, é encontrada quando $y = Ax$ é a projeção ortogonal do vetor L sobre $\mathcal{R}(A)$. Para isso basta escolher $V \in \mathcal{N}(A^T)$. Com essa escolha, o produto interno entre y e V será nulo para todo $x \in R^n$, logo:

$$\begin{aligned} \langle y, V \rangle &= \langle Ax, Ax - L \rangle \\ &= \langle x, A^T Ax - A^T L \rangle = 0, \end{aligned} \quad (3.6)$$

donde, resulta que a solução de mínimos quadrados satisfaz às equações normais:

$$A^T Ax = A^T L. \quad (3.7)$$

Sendo $r = n$, A tem o posto completo, $A^T A$ admite inversa $(A^T A)^{-1}$ e o vetor de parâmetros ajustados (x_a), pode ser representado explicitamente por:

$$x_a = (A^T A)^{-1} A^T L. \quad (3.8)$$

Se $r \neq n$, o sistema (3,7) sempre terá infinitas soluções.

O vetor dos resíduos é estimado como:

$$\hat{V} = Ax_a - L, \quad (3.9)$$

e o vetor das observações ajustadas será:

$$L_a = L + \hat{V}. \quad (3.10)$$

A representação do vetor ajustado (x_a) pela equação (3.8) tem um caráter apenas didático, pois a inversão direta de matrizes deve sempre que possível ser evitada, por causa do custo computacional dos algoritmos, e das perdas de informações numéricas devido às operações.

É comum, na prática, as observações serem realizadas por métodos diferentes, ou por equipamentos diferentes, ou ainda, serem correlacionadas entre si. O ajustamento também deve contemplar tal situação. Isto é feito, introduzindo uma matriz de pesos Q_L^{-1} para as observações. Essa matriz de pesos é simétrica e, em geral, definida positiva (ver seção 3.2.4).

O problema de mínimos quadrados linear, equação (3.5), para o caso ponderado pode ser definido como:

$$\min_{\mathbf{x} \in \mathbb{R}^n} \frac{1}{2} \|\mathbf{Ax} - \mathbf{L}\|_{Q_L^{-1}}^2, \quad (3.11)$$

onde:

$$\|\cdot\|_{Q_L^{-1}}^2 = (\cdot)^T Q_L^{-1} (\cdot) : \text{define uma norma de vetor}^1.$$

A matriz Q_L^{-1} sendo simétrica e definida positiva, admite a decomposição de Cholesky (ver seção 3.1.3.1):

$$Q_L^{-1} = \mathbf{W}^T \mathbf{W}, \quad (3.12)$$

onde:

\mathbf{W} : é o fator triangular superior de Cholesky da matriz Q_L^{-1} .

O problema (3.11), usando a decomposição (3.12), pode ser reescrito de uma forma equivalente, como:

$$\min_{\mathbf{x} \in \mathbb{R}^n} \frac{1}{2} \|\mathbf{W}(\mathbf{Ax} - \mathbf{L})\|_2^2. \quad (3.13)$$

Outra forma equivalente para o problema (3.13), seria definir a matriz \mathbf{W}

¹ Definições de normas de vetor e normas de matriz são encontradas em: NOBLE e DANIEL (1977,p.131-167), GOLUB e VAN LOAN (1989,p.53-60), GILL;MURRAY e WRIGHT (1991,p.21-27) e GARNÉS (1996, p. 7-9).

como a matriz raiz quadrada² de Q_L^{-1} , isto é:

$$W = U \Sigma U^T = Q_L^{-1/2}, \quad (3.14)$$

onde:

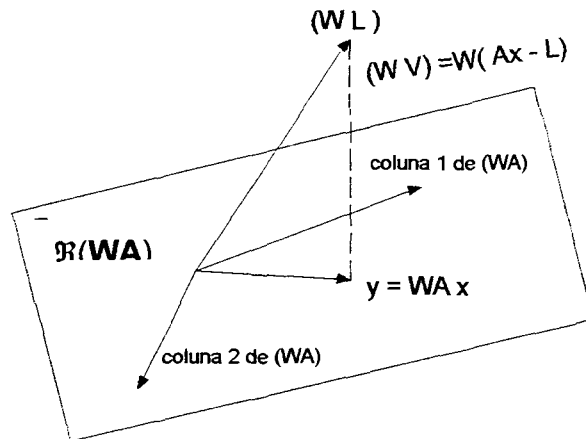
W : é uma matriz simétrica e definida positiva;

Σ : é uma matriz diagonal formada pela raiz quadrada dos autovalores de Q_L^{-1} ;

U : é uma matriz ortogonal formada dos autovetores normalizados, associados aos autovalores de Q_L^{-1} , na ordem em que aparecem em Σ .

A fig.3.2 ilustra a geometria do problema de mínimos quadrados linear ponderado para $A \in R^{3 \times 2}$.

FIGURA 3.2 – INTERPRETAÇÃO GEOMÉTRICA DO PROBLEMA DE MÍNIMOS QUADRADOS LINEAR PONDERADO



A solução é obtida quando:

$$\begin{aligned} \langle y, WV \rangle &= \langle WAx, WAx - WL \rangle \\ &= \langle x, A^T Q_L^{-1} Ax - A^T Q_L^{-1} L \rangle = 0, \quad \forall x \in R^n, \end{aligned} \quad (3.15)$$

² Propriedades da matriz raiz quadrada são encontradas em: JOHNSON e WICHERN (1988, p.51). Ver também a decomposição de valor singular, seção 3.1.3.3.

logo, a solução do problema de mínimos quadrados linear ponderado satisfaz o sistema de equações normais:

$$\mathbf{A}^T \mathbf{Q}_L^{-1} \mathbf{A} \mathbf{x} = \mathbf{A}^T \mathbf{Q}_L^{-1} \mathbf{L}. \quad (3.16)$$

A matriz \mathbf{A} tendo posto completo, $(\mathbf{A}^T \mathbf{Q}_L^{-1} \mathbf{A})$ admite inversa $(\mathbf{A}^T \mathbf{Q}_L^{-1} \mathbf{A})^{-1}$, e o vetor dos parâmetros ajustados (\mathbf{x}_a), pode ser representado por:

$$\mathbf{x}_a = (\mathbf{A}^T \mathbf{Q}_L^{-1} \mathbf{A})^{-1} \mathbf{A}^T \mathbf{Q}_L^{-1} \mathbf{L}. \quad (3.17)$$

Com \mathbf{x}_a estimado pela equação (3.17), calcula-se o vetor dos resíduos estimados ($\hat{\mathbf{V}}$), equação (3.9), e o vetor das observações ajustadas (\mathbf{L}_a), equação (3.10).

3.1.2 Mal-Condicionamento dos Sistemas de Equações Lineares

É muito comum que as observações GPS sejam realizadas a uma taxa de 1s, 5s e 15s. Com essas taxas, a variação da geometria dos satélites é muito pequena e as linhas da matriz \mathbf{A} (formada das derivadas parciais das equações de observação com relação aos parâmetros), de uma época para outra, são próximas de serem linearmente dependentes. A consequência disso, é que no processamento relativo com fase da portadora, o sistema de equações normais (3.16), possa ser mal-condicionado.

Reescrevendo o sistema de equações normais (3.16), como:

$$\mathbf{N} \mathbf{x} = \mathbf{u}, \quad (3.18)$$

vem que:

$$\mathbf{N} = \mathbf{A}^T \mathbf{Q}_L^{-1} \mathbf{A};$$

$$\mathbf{u} = \mathbf{A}^T \mathbf{Q}_L^{-1} \mathbf{L}.$$

Se uma pequena perturbação no vetor \mathbf{u} , por exemplo, $\delta \mathbf{u}$, e/ou uma pequena

perturbação na matriz \mathbf{N} , por exemplo, $\delta\mathbf{N}$, causar uma grande variação no vetor solução \mathbf{x} , por exemplo, $\|\delta\mathbf{x}\|_2 \gg \|\delta\mathbf{u}\|_2$, então, o sistema de equações (3.18) é um sistema mal-condicionado.

O número de condição³ generalizado de uma matriz é definido como o produto de sua norma pela norma de sua pseudo-inversa⁴ (GILL; MURRAY; WRIGHT, 1991, p. 171). Para a matriz \mathbf{N} , define-se:

$$C(\mathbf{N}) = \|\mathbf{N}\| \|\mathbf{N}^+\|, \quad (3.19)$$

onde:

$\|\cdot\|$: denota uma norma de matriz;

\mathbf{N}^+ : é a pseudo-inversa de \mathbf{N} .

O número de condição pode ser visto como um indicador sobre o condicionamento do sistema, pois ele informa sobre a máxima variação relativa que o vetor solução pode sofrer, quando perturbações no vetor \mathbf{u} ou na matriz \mathbf{N} ocorrerem. As expressões a seguir ilustram isto (GILL; MURRAY; WRIGHT, 1991, p. 197-198):

$$\frac{\|\delta\mathbf{x}\|}{\|\mathbf{x}\|} \leq C(\mathbf{N}) \frac{\|\delta\mathbf{u}\|}{\|\mathbf{u}\|} ; \quad (3.20)$$

$$\frac{\|\delta\mathbf{x}\|}{\|\mathbf{x} + \delta\mathbf{x}\|} \leq C(\mathbf{N}) \frac{\|\delta\mathbf{N}\|}{\|\mathbf{N}\|} \quad (3.21)$$

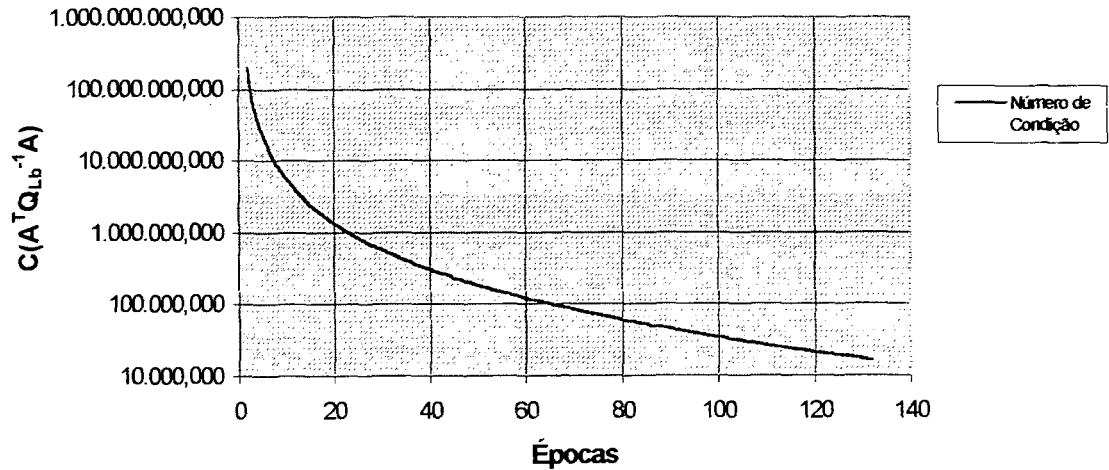
O gráfico apresentado na fig.3.3, foi preparado para ilustrar o comportamento do número de condição na norma-2 de matriz, do sistema de equações normais formado do ajustamento das duplas diferenças de fase (equação (2.93) sem os termos da refração troposférica e ionosférica), envolvendo 7 satélites rastreados a uma taxa de

³ LUGNANI (1975, p.65) e GEMAEL (1994, p.79) apresentam dois números de condição particulares, denominados, número de TURING e número de TODD.

⁴ A pseudo-inversa de uma matriz é a única inversa generalizada que satisfaz as quatro condições de Penrose (LAWSON; HANSON, 1974, p.38). Se \mathbf{N} é não-singular, então $\mathbf{N}^+ = \mathbf{N}^{-1}$.

coleta de dados de 15 s. O sistema de equações normais foi formado acumulando os dados seqüencialmente, época por época (ver ajustamento seqüencial, seção 3.1.4).

FIGURA 3.3 – NÚMERO DE CONDIÇÃO DA MATRIZ DE EQUAÇÕES NORMAIS DA DUPLA DIFERENÇA DE FASE



NOTA: experimento com linha de base de 185 m, 7 satélites rastreados a uma taxa de 15s – ajustamento seqüencial.

O gráfico (fig. 3.3) mostra que o número de condição $C(N)$, é bastante elevado para as primeiras épocas, a ordem de grandeza é de 1.10^8 .

Usando a matriz W , a partir das equações (3.12) ou (3.14), e com a equação (3.15), o sistema de equações normais pode ser representado por:

$$(WA)^T WA \mathbf{x} = (WA)^T W \mathbf{L} \quad , \quad (3.22)$$

onde:

$$N = (WA)^T WA.$$

A equação (3.19) pode se reescrita na forma:

$$C(N) = \|(WA)^T (WA)\| \|[(WA)^T (WA)]^+\| \quad . \quad (3.23)$$

Da definição da norma-2 de matriz, tem-se:

$$\|(\mathbf{WA})^T(\mathbf{WA})\|_2 = \lambda_{\max} ; \quad (3.24)$$

$$\|[(\mathbf{WA})^T(\mathbf{WA})]^+\|_2 = \frac{1}{\lambda_{\min}} ; \quad (3.25)$$

$$\|(\mathbf{WA})\|_2 = \sigma_{\max} = \sqrt{\lambda_{\max}} ; \quad (3.26)$$

$$\|(\mathbf{WA})^+\|_2 = \frac{1}{\sigma_{\min}} = \frac{1}{\sqrt{\lambda_{\min}}} , \quad (3.27)$$

onde:

$\lambda_{\max}, \lambda_{\min}$: maior e menor autovalor de \mathbf{N} , respectivamente;

$\sigma_{\max}, \sigma_{\min}$: maior e menor valor singular de (\mathbf{WA}) , respectivamente.

O número de condição da matriz (\mathbf{WA}) , de acordo com a definição é:

$$C(\mathbf{WA}) = \|\mathbf{WA}\| \|\mathbf{WA}^+\| . \quad (3.28)$$

Usando a norma-2 e as relações (3.24) à (3.27), obtém-se:

$$C(\mathbf{N}) = \frac{\lambda_{\max}}{\lambda_{\min}} = \left(\frac{\sigma_{\max}}{\sigma_{\min}} \right)^2 = C(\mathbf{WA})^2. \quad (3.29)$$

A análise anterior mostra que o número de condição da matriz de equações normais (na norma-2)⁵, é o quadrado do número de condição da matriz (\mathbf{WA}) . Esse resultado induz à utilização de métodos de solução do problema de mínimos quadrados linear, sem formação das equações normais.

⁵ A análise é semelhante para outras normas, a menos de uma constante, devido ao fato de que todas as normas em \mathbb{R}^n são equivalentes. Ver GOLUB e VAN LOAN (1989, p. 57).

Uma das formas de analisar o condicionamento do problema de mínimos quadrados, equação (3.13), sem a formação das equações normais, é (GILL; MURRAY; WRIGHT, 1991, p. 236-237):

$$\frac{\|\delta \mathbf{x}\|_2}{\|\mathbf{x}\|_2} \leq C(\mathbf{WA}) \frac{\|\delta(\mathbf{WL})_{\mathbf{R}}\|_2}{\|(\mathbf{WL})_{\mathbf{R}}\|_2}, \quad (3.30)$$

onde:

$(\mathbf{WL})_{\mathbf{R}}$: componente de (\mathbf{WL}) em $\mathcal{R}(\mathbf{WA})$;

$\delta(\mathbf{WL})_{\mathbf{R}}$: perturbação da componente de \mathbf{WL} em $\mathcal{R}(\mathbf{WA})$.

$$\frac{\|\delta \mathbf{x}\|_2}{\|\mathbf{x}\|_2} \leq 2 C(\mathbf{WA}) e_{\mathbf{R}} + 4 C(\mathbf{WA})^2 e_{\mathbf{N}} \frac{\|(\mathbf{WL})_{\mathbf{N}}\|_2}{\|(\mathbf{WL})_{\mathbf{R}}\|_2} + O(e_{\mathbf{N}}^2), \quad (3.31)$$

com:

$$e_{\mathbf{R}} = \frac{\|\mathbf{E}_{\mathbf{R}}\|_2}{\|\mathbf{WA}\|_2}; \quad e_{\mathbf{N}} = \frac{\|\mathbf{E}_{\mathbf{N}}\|_2}{\|\mathbf{WA}\|_2}; \quad \mathbf{E}_{\mathbf{R}} = \mathbf{M}^T \mathbf{E}; \quad \mathbf{E}_{\mathbf{N}} = \mathbf{K}^T \mathbf{E},$$

onde:

\mathbf{E} : é uma matriz de perturbação em \mathbf{WA} ;

\mathbf{M} : base ortonormal de $\mathcal{R}(\mathbf{WA})$;

\mathbf{K} : base ortonormal de $\mathcal{N}[(\mathbf{WA})^T]$.

A expressão (3.31) afirma: se não ocorrer perturbação no espaço nulo da matriz $(\mathbf{WA})^T$, o número de condição ao quadrado $(C(\mathbf{WA})^2)$ não afeta na solução do sistema $(\mathbf{WA})\mathbf{x} = \mathbf{WL}$.

As perturbações no vetor de observações ponderado (\mathbf{WL}) e/ou na matriz (\mathbf{WA}) , podem acontecer devido ao arredondamento (ou truncamento) para a precisão de máquina (computador) na representação de ponto flutuante⁶, durante as operações de cálculo.

⁶ Sobre aritmética de precisão finita e medidas de erro, consultar DENNIS JR e SCHNABEL (1983, p.10-12), GILL; MURRAY; e WRIGHT (1991, p. 33-39).

Quando esse for o caso, existe uma relação entre o número de condição da matriz de coeficientes (**WA**) e a unidade de arredondamento do computador (u), para garantir uma solução confiável do sistema de equações normais (3.16). Essa relação é (GILL; MURRAY; WRIGHT, 1991, p. 225), (BJÖRCK, 1991, p. 23):

$$C(\mathbf{WA}) < \frac{1}{\sqrt{u}}. \quad (3.32)$$

Para maiores detalhes quanto ao condicionamento do problema de mínimos quadrados linear, consultar as referências citadas nessa seção. Várias experiências numéricas sobre o assunto, estão relatadas em: GARNÉS (1996), GARNÉS; SAMPAIO e DALMOLIN (1997), GARNÉS e SAMPAIO (1999).

3.1.3 Métodos e Algoritmos para a Solução do Problema de Mínimos Quadrados Linear

De acordo com o que foi discutido nas duas seções anteriores, o problema de mínimos quadrado linear pode ser representado por⁷:

$$\min_{\mathbf{x} \in \mathbb{R}^n} \frac{1}{2} \|\mathbf{Ax} - \mathbf{L}\|_2^2, \quad (3.33)$$

onde:

$$\mathbf{A} = \mathbf{WA};$$

$$\mathbf{L} = \mathbf{WL};$$

$$\mathbf{Q}_L^{-1} = \mathbf{W}^T \mathbf{W}.$$

Obs.: Com essa substituição pode-se usar a norma-2 na equação (3.33). A substituição é apenas para facilitar a notação.

Lembrando que quando o problema não é ponderado, a matriz **W** é igual à

⁷ Deve ser observado que o resíduo estimado ($\hat{\mathbf{V}}$), equação (3.9), é calculado com a matriz original **A** e não com **A=WA**.

matriz identidade (**I**), isto é, $\mathbf{W} = \mathbf{I}$.

Uma observação importante vista na seção 3.1.2 é a relação (3.32). Ela mostra que devido a erros de arredondamento, só é confiável resolver o problema de mínimos quadrados linear via equações normais, se o número de condição da matriz $\mathbf{A}=\mathbf{W}\mathbf{A}$ for inferior ao inverso da raiz quadrada da unidade de arredondamento do computador.

Tendo em consideração esses detalhes, pode-se optar por um ou mais dos métodos apresentados na sequência.

3.1.3.1 Decomposição de Cholesky

Se \mathbf{N} é uma matriz simétrica e definida positiva⁸, então ela permite a decomposição de Cholesky:

$$\mathbf{N} = \mathbf{E} \mathbf{E}^T, \quad (3.34)$$

onde:

\mathbf{E} : é uma matriz triangular inferior.

Com a decomposição de Cholesky, a solução do sistema de equações normais $\mathbf{N}\mathbf{x} = \mathbf{u}$, se resume à solução de dois sistemas triangulares:

- resolver para \mathbf{y} o sistema triangular inferior $\mathbf{E}\mathbf{y} = \mathbf{u}$;
- resolver para \mathbf{x} o sistema triangular superior $\mathbf{E}^T\mathbf{x} = \mathbf{y}$.

Existem diferentes versões de algoritmos para a obtenção da decomposição de Cholesky. Alguns deles são encontrados em GASTINEL (1971)⁹, BJÖRCK (1991, p. 21), GOLUB e VAN LOAN (1989, p. 142-147).

A seguir é descrita a versão produto externo extraída de GOLUB e VAN

⁸ De acordo com MARTINEZ e SANTOS (1995, p. 43), a forma mais econômica para se averiguar se uma matriz simétrica é definida positiva, é tentar fazer a sua fatoração de Cholesky.

⁹ A versão apresentada em GASTINEL (1971) também é descrita em GARNÉS (1996, p. 84). Nesta subrotina estão incluídas as soluções dos sistemas $\mathbf{E}\mathbf{y} = \mathbf{u}$ e $\mathbf{E}^T\mathbf{x} = \mathbf{y}$.

LOAN (1989, p.143). A apresentação do algoritmo segue aproximadamente a forma da linguagem de programação do MATLAB.

Algoritmo 3.1 (Decomposição de Cholesky: versão produto externo). Dada uma matriz simétrica e definida positiva $N \in R^{n \times n}$, o seguinte algoritmo calcula uma matriz triangular inferior E tal que $N = EE^T$. Para todo $i \geq j$, $E(i,j)$ sobrescreve $N(i,j)$.

```

for k = 1: n
    N(k,k) = sqrt(N(k,k))
    if N(k,k) ≠ 0
        for i = k+1 : n ,
            N(i,k) = N(i,k) / N(k,k);
        end
        for j = k+1 : n
            for i = j : n ,
                N(i,j) = N(i,j) - N(i,k) N(j,k);
            end;
        end;
    end;
end
end

```

Para a determinação da matriz de variância-covariância dos parâmetros ajustados (ver seção 3.2), será necessário à inversão da matriz N .

Da equação (3.34), usando uma das propriedades da inversão de matrizes, tem-se:

$$N^{-1} = (EE^T)^{-1} = E^{-T}E^{-1}. \quad (3.35)$$

Na inversão da matriz triangular E , cada coluna de E^{-1} é encontrada resolvendo-se o sistema triangular:

$$Ex_i = I_i, \quad i = 1, \dots, n, \quad (3.36)$$

onde:

x_i : é a i -ésima coluna de E^{-1} ;

I_i : é a i -ésima coluna da matriz identidade.

O algoritmo a seguir, já otimizado, extraído de JONGE e TIBERIUS (1996, p.43), pode ser usado para obter a inversa da matriz E .

Algoritmo 3.2 (Inversão de matriz triangular inferior) : Dada uma matriz triangular inferior $E \in R^{n \times n}$, o algoritmo calcula a inversa $X = E^{-1}$. A relação entre a matriz E e sua inversa X é $EX = I$.

```

for i = 1: n
  for j = 1: i -1
    X(i, j) = -dot(E(i, j : i -1), X(j : i -1, j))/ E(i, i);
  end
  X(i, i) = 1/ E(i, i);
end

```

A notação usada no algoritmo está na linguagem do MATLAB. A função *dot* representa o produto interno.

3.1.3.2 Decomposição QR

Usando transformações de Householder, Gram-Shimidt ou de Givens, a matriz $A \in R^{m \times n}$, com posto completo, pode ser decomposta em $A = QR$, onde $Q \in R^{m \times m}$ é uma matriz ortogonal e $R \in R^{m \times n}$ é triangular superior à direita.

Fazendo a substituição de A por QR no problema (3.33), obtém-se:

$$\min_{x \in R^n} \frac{1}{2} \| Ax - L \|_2^2 = \min_{x \in R^n} \frac{1}{2} \{ \| R_{(n)} x - (Q^T L)_{(n)} \|_2^2 + \| (Q^T L)_{(m-n)} \|_2^2 \}, \quad (3.37)$$

o que é equivalente a:

$$\min_{x \in R^n} \frac{1}{2} \| R_{(n)} x - (Q^T L)_{(n)} \|_2^2, \quad (3.38)$$

e para a solução basta resolver o sistema triangular superior:

$$R_{(n)} x = (Q^T L)_{(n)}. \quad (3.39)$$

Obs.: o (n) subscrito representa as n primeiras linhas da matriz R e do vetor $Q^T L$.

O algoritmo a seguir para obtenção da decomposição QR , é uma adaptação da versão apresentada em DENNIS JR e SCHNABEL (1983, p.304-307).

Algoritmo 3.3 (Decomposição QR). Dada uma matriz $A \in R^{m \times n}$, com $m > n$ e $\text{posto}(A) = n$, o seguinte algoritmo calcula uma matriz ortogonal $Q \in R^{m \times m}$ e uma triangular superior a direita $R \in R^{m \times n}$ tal que $A = QR$. Para todo $i \geq j$, $R(i,j)$ sobrescreve $A(i,j)$.

```

for k=1: n
    η = abs(A(k,k))
    for i = k : m
        if abs(A(i,k)) > η , η= abs(A(i,k)); end; end
    if η = 0 % a matriz A é singular
        m1(k) = 0 ; m2(k) = 0
    else
        for i = k : m, A(i,k) = A(i,k)/ η ; end
        s = 0
        for i = k : m, s = s + A(i,k) A(i,k); end
        if A(k,k) > 0 % determinação do sinal de A(k,k)
            sign = 1; elseif A(k,k) < 0, sign = -1 ; else, sign = 0; end
        σ = sign sqrt(s); A(k,k) = A(k,k) + σ ;
        m1(k) = σ A(k,k) ; m2(k) = -η σ
        for j = k+1 : n
            s = 0
            for i = k : m, s = s + A(i,k) A(i,j); end
            τ = s / m1(k)
            for i = k : m, A(i,j) = A(i,j) - τ A(i,k); end; end; end; end
        for i = 1 : m % formando a matriz identidade
            for j = 1 : m
                if i = j , Q(i,j) = 1; else, Q(i,j) = 0; end; end; end
        for k = 1 : n % formando a matriz QT
            if m1(k) ≠ 0
                for j = 1 : m
                    τ = 0
                    for i = k : m, τ = τ + A(i,k) Q(i,j); end
                    τ = τ / m1(k)
                    for i = k : m, Q(i,j) = Q(i,j) - τ A(i,k) ; end; end; end; end
        for i = 1 : m % formando a matriz Q
            for j = i : m
                if i≠j
                    s = Q(i,j) ; Q(i,j) = Q(j,i) ; Q(j,i) = s; end; end; end
        for i = 1 : n % Completando R na parte triangular superior de A
            A(i,i) = m2(i); end

```

Obs.: o algoritmo acima calcula toda a matriz Q , mas para a solução do sistema (3.37) não é necessário formá-la explicitamente, toma o método mais lento. Ver a implementação em DENNIS JR e SCHNABEL (1983, p.304-307).

O método de solução do problema de mínimos quadrados linear usando a decomposição QR é muito estável. A descrição detalhada do método pode ser encontrada em GARNÉS (1996, p. 85-92).

A inversa de N depois da realização da decomposição QR de A é obtida como:

$$N^{-1} = (A^T A)^{-1} = (R^T Q^T Q R)^{-1} = (R^T R)^{-1} = R_{(n)}^{-1} R_{(n)}^{-T}. \quad (3.40)$$

A inversa da triangular superior $R_{(n)}$ é obtida de uma forma semelhante à inversa da triangular inferior E . O algoritmo 3.2 pode ser adaptado para fazer isto.

3.1.3.3 Decomposição de valor singular - SVD

Dada qualquer matriz $A \in R^{m \times n}$, chama-se decomposição de valor singular de A , a seguinte decomposição:

$$A = U \Sigma V^T, \quad (3.41)$$

onde:

$U \in R^{m \times m}$: é uma matriz ortogonal;

$\Sigma \in R^{m \times n}$ tem a forma

$$\Sigma = \begin{bmatrix} \Sigma_r & 0 \\ 0 & 0 \end{bmatrix},$$

com: $\Sigma_r = \text{diag}(\sigma_1, \sigma_2, \dots, \sigma_r)$, $\sigma_i > 0$, $i = 1, \dots, r$,

onde: $r = \text{posto}(A)$; e σ_i são os valores singulares de A ;

$V \in R^{n \times n}$: é uma matriz ortogonal.

A prova da existência da decomposição de valor singular é encontrada em LAWSON e HANSON (1974, p.18), GOLUB e VAN LOAN (1989, p. 71). A descrição detalhada do método é encontrada em GOLUB e REINCH (1970),

LAWSON e HANSON (1974, p. 107-118) e GARNÉS (1996, p. 92-108).

Fazendo a substituição $\mathbf{A} = \mathbf{U}\Sigma\mathbf{V}^T$, com $\text{posto}(\mathbf{A})=n$, no problema de mínimos quadrados linear, obtém-se:

$$\min_{\mathbf{x} \in R^n} \frac{1}{2} \|\mathbf{Ax} - \mathbf{L}\|_2^2 = \min_{\mathbf{x} \in R^n} \frac{1}{2} \{ \|\Sigma_{(n)} \mathbf{V}^T \mathbf{x} - (\mathbf{U}^T \mathbf{L})_{(n)}\|_2^2 + \|(\mathbf{U}^T \mathbf{L})_{(m-n)}\|_2^2 \}. \quad (3.42)$$

A minimização de (3.42) ocorre no mínimo de:

$$\min_{\mathbf{x} \in R^n} \frac{1}{2} \{ \|\Sigma_{(n)} \mathbf{V}^T \mathbf{x} - (\mathbf{U}^T \mathbf{L})_{(n)}\|_2^2 \}, \quad (3.43)$$

sendo alcançado em:

$$\mathbf{x}_a = \mathbf{V}\Sigma_{(n)}^{-1}(\mathbf{U}^T \mathbf{L})_{(n)}. \quad (3.44)$$

Obs.: o (n) subscrito indica as n primeiras linhas da matriz Σ e do vetor $(\mathbf{U}^T \mathbf{L})$.

Denotando Σ^+ a matriz $n \times m$ cujos elementos não nulos da diagonal principal são o inverso dos valores singulares, ou seja, $\frac{1}{\sigma_1}, \frac{1}{\sigma_2}, \dots, \frac{1}{\sigma_n}$, a expressão (3.44) pode ser representada por:

$$\mathbf{x}_a = \mathbf{V}\Sigma^+ \mathbf{U}^T \mathbf{L}. \quad (3.45)$$

A matriz $\mathbf{V}\Sigma^+ \mathbf{U}^T \in R^{n \times m}$ é a pseudo-inversa da matriz \mathbf{A} e geralmente é denotada por \mathbf{A}^+ .

Se a matriz \mathbf{A} tem deficiência de posto, por exemplo, $\text{posto}(\mathbf{A})=r < n$, então os elementos não nulos de Σ^+ serão $\frac{1}{\sigma_1}, \frac{1}{\sigma_2}, \dots, \frac{1}{\sigma_r}, 0, \dots, 0$, e a solução pela pseudo-inversa, será:

$$\mathbf{x}_a^+ = \mathbf{A}^+ \mathbf{L}, \quad (3.46)$$

a qual corresponde a solução de comprimento mínimo, isto é, $\|\mathbf{x}_a^+\|_2 = \text{mínimo}$. Essa solução é muito usada em Geodésia no ajustamento livre (GARNÉS e DALMOLIN, 1998), GEMAEL (1994, p.248).

A inversa da matriz \mathbf{N} , usando a decomposição de valor singular é:

$$\mathbf{N}^{-1} = (\mathbf{A}^T \mathbf{A})^{-1} = (\mathbf{V} \Sigma^T \mathbf{U}^T \mathbf{U} \Sigma \mathbf{V}^T)^{-1} = (\mathbf{V} \Sigma^T \Sigma \mathbf{V}^T)^{-1} = \mathbf{V} \mathbf{D}^{-1} \mathbf{V}^T, \quad (3.47)$$

onde:

$$\mathbf{D} = \Sigma^T \Sigma = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n), \text{ sendo } \lambda_i, i = 1, \dots, n \text{ os autovalores de } \mathbf{A}^T \mathbf{A}.$$

A decomposição de valor singular é uma das ferramentas mais importantes na análise do problema de mínimos quadrados linear, pois por trabalhar sem formar explicitamente o sistema de equações normais, foge do condicionamento quadrado. Além disso, fornece: os números de condição das matrizes \mathbf{A} e \mathbf{N} ; uma expressão imediata para a matriz variância-covariância dos parâmetros ajustados com base na equação (3.47); e as bases ortogonais dos 4 subespaços fundamentais da matriz \mathbf{A} (ver tab.3.1). É um dos principais métodos para determinar o posto de uma matriz, e além de tudo, fornece ainda as orientações e dimensões dos semi-eixos do elipsóide n -dimensional dos erros, por meio da equação (3.47).

A subrotina da decomposição de valor singular pode ser encontrada em PRESS et al.,(1986, p.60-64) ou, também, em GARNÉS (1996, p.108-112).

3.1.3.4 Método dos gradientes conjugados

Os métodos baseados nas decomposições de Cholesky, QR e decomposição de valor singular (SVD) são métodos diretos de solução do problema de mínimos quadrados linear. Dentre os métodos diretos atualmente existentes, são os mais recomendados. No entanto, uma vez iniciado o processo de cálculo, seja com precisão computacional simples ou dupla, só resta esperar a solução com a precisão dada pelo método.

Os métodos iterativos, ao contrário dos métodos diretos, têm a opção de imposição da condição de parada. Isto é bastante útil quando não é necessária tanta precisão numérica (precisão computacional dupla), pois se pode economizar tempo computacional nos cálculos. Também pode ser útil no caso contrário, isto é, quando a

precisão computacional é mais importante do que o tempo. Desta forma, os métodos iterativos permitem um controle computacional sobre a precisão e o tempo para solução.

Dentre os métodos iterativos para a resolução de sistemas lineares, os mais populares são os métodos de Jacobi, Gauss-Seidel, sobre-relaxação sucessiva–SOR, e gradientes conjugados. Contudo, experiências realizadas por GARNÉS e SAMPAIO (1999, p. 226) no ajustamento de uma linha de base com as duplas diferenças de fase GPS, envolvendo apenas 5 épocas de observações a uma taxa de 30s, mostraram que entre os métodos SOR, Gauss-Seidel e gradientes conjugados, apenas o método dos gradientes conjugados converge com sucesso.

Descrição do método

Dado um ponto inicial \mathbf{x}_0 , o método dos gradientes conjugados iterativamente busca novos pontos \mathbf{x}_k , $k=1,2, \dots$, até alcançar o ponto \mathbf{x}_a minimizador da forma quadrática:

$$f(\mathbf{x}) = \frac{1}{2} \mathbf{x}^T \mathbf{N} \mathbf{x} - \mathbf{x}^T \mathbf{u} + \frac{1}{2} \mathbf{L}^T \mathbf{L}, \quad (3.48)$$

onde:

$\mathbf{N} = \mathbf{A}^T \mathbf{A}$: matriz das equações normais, simétrica e definida positiva;

$\mathbf{u} = \mathbf{A}^T \mathbf{L}$: vetor das equações normais.

O vetor gradiente e a matriz Hessiana de $f(\mathbf{x})$ são, respectivamente:

$$\nabla f(\mathbf{x}) = \mathbf{N} \mathbf{x} - \mathbf{u}; \quad (3.49)$$

$$\nabla^2 f(\mathbf{x}) = \mathbf{N}. \quad (3.50)$$

Se \mathbf{x}_a é minimizador da forma quadrática de mínimos quadrados (3.48), ele anula o vetor gradiente $\nabla f(\mathbf{x}_a) = \mathbf{0}$. Desta forma, o ponto \mathbf{x}_a , é também solução do sistema de equações normais:

$$\mathbf{N}\mathbf{x}_a = \mathbf{u} . \quad (3.51)$$

Denotando por \mathbf{p}_k as direções de descida em $f(\mathbf{x})$, o método dos gradientes conjugados fica caracterizado pela escolha de direções \mathbf{p}_k 's mutuamente conjugadas, isto é, $\mathbf{p}_{k-1}^T \mathbf{N} \mathbf{p}_k = 0$, $k=2,3, \dots$.

A definição de \mathbf{p}_k N-conjugado com \mathbf{p}_{k-1} é feita mediante a determinação do parâmetro β_k da seguinte combinação linear:

$$\mathbf{p}_k = \mathbf{r}_{k-1} + \beta_k \mathbf{p}_{k-1}, \quad k = 2,3,4, \dots \quad (3.52)$$

onde:

$$\mathbf{r}_{k-1} = -\nabla f(\mathbf{x}_{k-1}), \text{ vale para } k = 1,2, \dots ;$$

$$\mathbf{p}_1 = \mathbf{r}_0;$$

$$\mathbf{x}_0 = \mathbf{0} \text{ (como é geralmente escolhido).}$$

A expressão de β_k para esse propósito é:

$$\beta_k = \frac{\mathbf{r}_k^T \mathbf{r}_k}{\mathbf{r}_{k-1}^T \mathbf{r}_{k-1}}, \quad k = 2,3,4, \dots \quad (3.53)$$

Uma vez definida a direção de descida \mathbf{p}_k no ponto \mathbf{x}_{k-1} , o valor de \mathbf{x}_k é alcançado pela minimização unidimensional de (3.48):

$$\min_{\alpha_k \in \mathbb{R}} f(\mathbf{x}_{k-1} + \alpha_k \mathbf{p}_k) . \quad (3.54)$$

O resultado desta minimização é:

$$\alpha_k = \frac{\mathbf{r}_{k-1}^T \mathbf{r}_{k-1}}{\mathbf{p}_k^T \mathbf{N} \mathbf{p}_k}, \quad (3.55)$$

e com esse valor calcula-se o novo ponto:

$$\mathbf{x}_k = \mathbf{x}_{k-1} + \alpha_k \mathbf{p}_k . \quad (3.56)$$

A descrição anterior torna-se mais esclarecida pelo algoritmo 3.4 (relatado a seguir), retirado a partir de GOLUB e VAN LOAN, (1989, p. 524).

Algoritmo 3.4 (Gradientes conjugados) . Se $N \in R^{n \times n}$ é simétrica e definida positiva e $u \in R^n$, então o seguinte algoritmo calcula $x \in R^n$ de modo que $Nx = u$.

```

k= 0;  $x_0 = 0$ ;  $r_0 = u$ ;  $\rho_0 = \|r_0\|_2^2$ 
while  $\sqrt{\rho_k} > tol$  or  $k < k_{max}$ 
    k=k+1
    if k=1
         $p_1 = r_0$ 
    else
         $\beta_k = r_{k-1}^T r_{k-1} / r_{k-2}^T r_{k-2}$ 
         $p_k = r_{k-1} + \beta_k p_{k-1}$ 
    end
     $w_k = N p_k$ 
     $\alpha_k = \rho_{k-1} / p_k^T w_k$ 
     $x_k = x_{k-1} + \alpha_k p_k$ 
     $r_k = r_{k-1} - \alpha_k w_k$ 
     $\rho_k = \|r_k\|_2^2$ 
end

```

Teoricamente o método dos gradientes conjugados deve convergir para a solução (x_a) da quadrática (3.48), em no máximo $k = n$ iterações¹⁰. No entanto, na prática, por causa de erros de arredondamento, as direções de descida (p_k) podem não ser totalmente N-conjugadas, tornando-se necessárias mais do que n iterações. A convergência do método também pode ser expressa com base no condicionamento do sistema pelo seguinte resultado (GOLUB e VAN LOAN, 1989, p. 525), (LUENBERG, 1973, p. 187), (YUAN, 1993, p.22):

$$\|x_a - x_k\|_N \leq 2 \|x_a - x_0\|_N \left(\frac{\sqrt{C(N)} - 1}{\sqrt{C(N)} + 1} \right)^k, \quad (3.57)$$

¹⁰ Os teoremas provando a convergência do método dos gradientes conjugados podem ser encontrados em MARTINEZ e SANTOS (1995, p. 45-54), GOLUB e VAN LOAN (1989, p. 516-526).

onde:

$$\|\cdot\|_N = \sqrt{(\cdot)^T N (\cdot)} ;$$

$C(N)$: número de condição de N na norma-2;

k : número da iteração;

\mathbf{x}_a : ponto de mínimo da quadrática (3.48);

\mathbf{x}_0 : ponto inicial do processo iterativo.

Se o sistema for mal-condicionado, a convergência pode ser extremamente lenta. A tabela 3.2 mostra um comparativo da variação do fator $2 \left(\frac{\sqrt{C(N)} - 1}{\sqrt{C(N)} + 1} \right)^k$ para um sistema onde o número de condição $C(N) = 100$ e, outro, onde o número de condição $C(N) = 1.10^8$ (ver fig. 3.3).

TABELA 3.2 – CONVERGÊNCIA DO MÉTODO DOS GRADIENTES CONJUGADOS

ITERAÇÃO (K)	$C(N) = 100$	$C(N) = 1.10^8$
	$2 \left(\frac{\sqrt{C(N)} - 1}{\sqrt{C(N)} + 1} \right)^k$	$2 \left(\frac{\sqrt{C(N)} - 1}{\sqrt{C(N)} + 1} \right)^k$
1	1,636	2,000
2	1,339	1,999
3	1,095	1,999
4	0,896	1,998
5	0,733	1,998
.	.	.
.	.	.
.	.	.

Uma maneira de melhorar um pouco a razão de convergência do método quando aplicado ao problema de mínimos quadrados linear, é não formar explicitamente as equações normais. Para isto deve-se trabalhar com a forma fatorada $\mathbf{A}^T(\mathbf{L}-\mathbf{Ax}) = \mathbf{0}$. O algoritmo 3.5 (relatado a seguir), adaptado a partir de BJÖRCK (1991, p. 104), mostra como proceder.

Algoritmo 3.5 (Gradientes conjugados no problema de mínimos quadrados linear).
 Dado $A \in R^{m \times n}$, $m > n$, $\text{posto}(A) = n$ e $L \in R^m$, o algoritmo calcula $x \in R^n$ de modo que $A^T(L - Ax) = 0$.

```

k= 0; x0 = 0 ; r0 = L; s0=ATr0 ; γ0 = ||s0||22
while √γk > tol or k < kmax
  k=k+1
  if k=1
    pk = s0
  else
    βk = γk-1 / γk-2
    pk = sk-1 + βk pk-1
  end
  qk = Apk
  αk = γk-1 / ||qk||22
  xk = xk-1 + αk pk
  rk = rk-1 - αk qk
  sk = ATrk
  γk = ||sk||22
end

```

Segundo BJÖRCK (1991, p. 104), ELFVING¹¹, testou várias outras variantes do método dos gradientes conjugados, mas o melhor resultado encontrado foi com a versão acima.

Mesmo trabalhando com a forma fatorada $A^T(L - Ax) = 0$, na presença de problemas mal-condicionados, o método ainda é lento. Todavia, nestes casos, para tornar a convergência rápida, até em torno de n iterações, é possível realizar um condicionamento através de uma matriz $M \in R^{n \times n}$ não-singular.

Com a matriz de condicionamento, o problema de mínimos quadrados linear pode ser colocado na forma equivalente:

$$\min_{y \in R^n} \|(AM^{-1})y - L\|_2^2, \quad (3.58)$$

¹¹ ELFVING, T. Some numerical results obtained with two gradient methods for solving the linear least squares problem. **Technical Report** Li TH – MAT – R – 75-5. Department of Math. Linköping University, 1978.

com:

$$\mathbf{M}\mathbf{x} = \mathbf{y}.$$

O algoritmo 3.6, adaptado a partir de BJÖRCK (1991, p.109-110), mostra como resolver o problema (3.58) utilizando o método dos gradientes conjugados com a matriz de condicionamento.

Algoritmo 3.6 (Gradientes conjugados com condicionamento no problema de mínimos quadrados linear). Dado $\mathbf{A} \in \mathbb{R}^{m \times n}$, $m > n$, $\text{posto}(\mathbf{A}) = n$ e $\mathbf{L} \in \mathbb{R}^m$ e $\mathbf{M} \in \mathbb{R}^{n \times n}$ não-singular, o algoritmo calcula $\mathbf{x} \in \mathbb{R}^n$ de modo que $\mathbf{A}^T(\mathbf{L} - \mathbf{A}\mathbf{x}) = \mathbf{0}$.

```

k= 0;  $\mathbf{x}_0 = \mathbf{0}$  ;  $\mathbf{r}_0 = \mathbf{L}$ ;  $\mathbf{s}_0 = \mathbf{M}^{-T} \mathbf{A}^T \mathbf{r}_0$  ;  $\gamma_0 = \|\mathbf{s}_0\|_2^2$ 
while  $\sqrt{\gamma_k} > tol$  or  $k < k_{\max}$ 
    k=k+1
    if k=1
         $\mathbf{p}_k = \mathbf{s}_0$ 
    else
         $\beta_k = \gamma_{k-1} / \gamma_{k-2}$ 
         $\mathbf{p}_k = \mathbf{s}_{k-1} + \beta_k \mathbf{p}_{k-1}$ 
    end
     $\mathbf{q}_k = \mathbf{A} \mathbf{M}^{-1} \mathbf{p}_k$ 
     $\alpha_k = \gamma_{k-1} / \|\mathbf{q}_k\|_2^2$ 
     $\mathbf{x}_k = \mathbf{x}_{k-1} + \alpha_k \mathbf{M}^{-1} \mathbf{p}_k$ 
     $\mathbf{r}_k = \mathbf{r}_{k-1} - \alpha_k \mathbf{q}_k$ 
     $\mathbf{s}_k = \mathbf{M}^{-T} \mathbf{A}^T \mathbf{r}_k$ 
     $\gamma_k = \|\mathbf{s}_k\|_2^2$ 
end

```

Na literatura são encontradas muitas formas de se obter condicionadores, veja, por exemplo: BJÖRCK (1991), GOLUB e VAN LOAN (1989), SILVA (1997) e YUAN (1993). Mas por causa da estrutura da matriz \mathbf{A} no ajustamento com as duplas diferenças de fase, é melhor utilizar os condicionadores versão bloco com a decomposição QR.

Particionando-se a matriz \mathbf{A} como $\mathbf{A} = [\mathbf{A}_1 \ \mathbf{A}_2]$, onde $\mathbf{A}_1 \in \mathbb{R}^{m \times n_1}$ e $\mathbf{A}_2 \in \mathbb{R}^{m \times n_2}$ com $n_1 + n_2 = n$, realiza-se a decomposição QR numa forma econômica, separadamente para \mathbf{A}_1 e \mathbf{A}_2 ; $\mathbf{A}_1 = \mathbf{Q}_1 \mathbf{R}_1$, com $\mathbf{Q}_1 \in \mathbb{R}^{m \times n_1}$ e $\mathbf{R}_1 \in \mathbb{R}^{n_1 \times n_1}$; e, $\mathbf{A}_2 = \mathbf{Q}_2 \mathbf{R}_2$, com

$$\mathbf{Q}_2 \in R^{m \times n_2} \text{ e } \mathbf{R}_2 \in R^{n_2 \times n_2}.$$

O preconditionador \mathbf{M} terá a forma:

$$\mathbf{M} = \begin{bmatrix} \mathbf{R}_1 & 0 \\ 0 & \mathbf{R}_2 \end{bmatrix}; \quad \text{e} \quad \mathbf{M}^{-1} = \begin{bmatrix} \mathbf{R}_1^{-1} & 0 \\ 0 & \mathbf{R}_2^{-1} \end{bmatrix}. \quad (3.59)$$

O produto \mathbf{AM}^{-1} que aparece no algoritmo (3.6) já foi calculado, pois, $\mathbf{AM}^{-1} = [\mathbf{Q}_1 \ \mathbf{Q}_2]$.

Obs: as matrizes \mathbf{R}_1 e \mathbf{R}_2 também podem ser obtidas usando a decomposição de Cholesky de $\mathbf{A}_1^T \mathbf{A}_1$ e $\mathbf{A}_2^T \mathbf{A}_2$, respectivamente.

A inversa da matriz \mathbf{N} , para a determinação da matriz variância-covariância dos parâmetros ajustados deverá ser realizada pelo algoritmo 3.4. Para isto basta resolver n sistemas de equações, um para cada coluna de \mathbf{N}^{-1} , da forma:

$$\mathbf{N}\mathbf{x}_i = \mathbf{I}_i, \quad i = 1, \dots, n, \quad (3.60)$$

onde:

\mathbf{x}_i : i -ésima coluna da matriz \mathbf{N}^{-1} ;

\mathbf{I}_i : i -ésima coluna da matriz identidade.

O controle da precisão de convergência para as colunas da matriz inversa se faz mediante o parâmetro *tol* no algoritmo.

3.1.4 Ajustamento Sequencial

Se as observações são realizadas em épocas diferentes objetivando estimar um conjunto único de parâmetros, o ajustamento poderá ser realizado com todas as observações juntas, ou poderá ser realizado em blocos, acrescentando sequencialmente, a influência de cada um dos blocos de observações.

Quando o modelo é não-linear, deve-se primeiro aplicar um dos métodos de

ajustamento do problema de mínimos quadrados não-linear (seção 3.1.7), para algumas das observações apenas, e encontrar a solução \mathbf{x}_a . Essa solução por estar muito próxima da solução geral, deve ser considerada como ponto aproximado para a linearização de todos os blocos, inclusive os usados para obtê-la.

A partir do problema (3.13) e fig. 3.2, dois blocos de observações podem ser representados por:

$$\text{Bloco 1:} \quad \mathbf{W}_1 \mathbf{V}_1 = \mathbf{W}_1 \mathbf{A}_1 \mathbf{x} - \mathbf{W}_1 \mathbf{L}_1, \quad (3.61)$$

$$\text{Bloco 2:} \quad \mathbf{W}_2 \mathbf{V}_2 = \mathbf{W}_2 \mathbf{A}_2 \mathbf{x} - \mathbf{W}_2 \mathbf{L}_2. \quad (3.62)$$

As matrizes de pesos $\mathbf{W}_1^T \mathbf{W}_1 = \mathbf{Q}_1^{-1}$ e $\mathbf{W}_2^T \mathbf{W}_2 = \mathbf{Q}_2^{-1}$ são consideradas independentes entre os blocos.

Reescrevendo (3.61) e (3.62) numa única expressão matricial, tem-se:

$$\begin{bmatrix} \mathbf{W}_1 \mathbf{V}_1 \\ \mathbf{W}_2 \mathbf{V}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{W}_1 \mathbf{A}_1 \\ \mathbf{W}_2 \mathbf{A}_2 \end{bmatrix} \mathbf{x} - \begin{bmatrix} \mathbf{W}_1 \mathbf{L}_1 \\ \mathbf{W}_2 \mathbf{L}_2 \end{bmatrix}. \quad (3.63)$$

Formando o sistema de equações normais

$$\begin{bmatrix} \mathbf{A}_1^T \mathbf{W}_1^T & \mathbf{A}_2^T \mathbf{W}_2^T \end{bmatrix} \begin{bmatrix} \mathbf{W}_1 \mathbf{A}_1 \\ \mathbf{W}_2 \mathbf{A}_2 \end{bmatrix} \mathbf{x} = \begin{bmatrix} \mathbf{A}_1^T \mathbf{W}_1^T & \mathbf{A}_2^T \mathbf{W}_2^T \end{bmatrix} \begin{bmatrix} \mathbf{W}_1 \mathbf{L}_1 \\ \mathbf{W}_2 \mathbf{L}_2 \end{bmatrix} \quad (3.64)$$

resulta:

$$[\mathbf{N}_1 + \mathbf{N}_2] \mathbf{x} = [\mathbf{u}_1 + \mathbf{u}_2]. \quad (3.65)$$

A expressão (3.65) quando generalizada para t blocos, torna-se:

$$\left[\sum_{i=1}^t \mathbf{N}_i \right] \mathbf{x} = \left[\sum_{i=1}^t \mathbf{u}_i \right], \quad (3.66)$$

onde:

$$\mathbf{N}_i = \mathbf{A}_i^T \mathbf{Q}_i^{-1} \mathbf{A}_i ;$$

$$\mathbf{u}_i = \mathbf{A}_i^T \mathbf{Q}_i^{-1} \mathbf{L}_i.$$

GEMAEL (1994, p. 191-194) apresenta o ajustamento seqüencial colocando, além dessas, as expressões para obter a influência das novas observações nos parâmetros e, também, na matriz variância-covariância.

Para resolver o sistema (3.66), se o número de condição $C(\mathbf{N}_i)$ for “moderado” (ver equação (3.32)), pode-se usar qualquer um dos métodos descritos na seção (3.1.3). Mas, se for “elevado” (exemplo: 1.10^8) é melhor agrupar todas as observações e optar por métodos mais estáveis que trabalham com o sistema $\mathbf{Ax} = \mathbf{L}$, por exemplo: o método QR, decomposição de valor singular ou gradientes conjugados com condicionamento.

3.1.5 Problema de Mínimos Quadrados Inteiro

No caso do ajustamento GPS das observações da fase da portadora, aparecem como parâmetros as ambigüidades. Elas são valores de natureza inteira e quando estimadas corretamente fazem aumentar a precisão do posicionamento. O capítulo 4 é dedicado exclusivamente para tratar com a resolução das ambigüidades. Entretanto, as considerações iniciais são abordadas nessa e na próxima seção.

O vetor dos parâmetros \mathbf{x} , considerando as componentes reais e as componentes inteiras, pode ser colocado na forma particionada:

$$\mathbf{x} = \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{bmatrix}, \quad (3.67)$$

onde:

$\mathbf{x}_1 \in \mathbf{R}^{n_1}$: vetor de componentes reais;

$\mathbf{x}_2 \in \mathbf{Z}^{n_2}$: vetor de componentes inteiros.

O resíduo a partir do problema (3.13), com o vetor \mathbf{x} particionado na forma (3.67), pode ser representado por:

$$\mathbf{WV} = \mathbf{W}[\mathbf{A}_1 \ \mathbf{A}_2] \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{bmatrix} - \mathbf{WL}. \quad (3.68)$$

O sistema de equações normais para (3.68) torna-se:

$$\begin{bmatrix} \mathbf{A}_1^T \mathbf{Q}_L^{-1} \mathbf{A}_1 & \mathbf{A}_1^T \mathbf{Q}_L^{-1} \mathbf{A}_2 \\ \mathbf{A}_2^T \mathbf{Q}_L^{-1} \mathbf{A}_1 & \mathbf{A}_2^T \mathbf{Q}_L^{-1} \mathbf{A}_2 \end{bmatrix} \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{A}_1^T \mathbf{Q}_L^{-1} \mathbf{L} \\ \mathbf{A}_2^T \mathbf{Q}_L^{-1} \mathbf{L} \end{bmatrix}, \text{ ou} \quad (3.69)$$

$$\underbrace{\begin{bmatrix} \mathbf{N}_{11} & \mathbf{N}_{12} \\ \mathbf{N}_{21} & \mathbf{N}_{22} \end{bmatrix}}_{\mathbf{N}} \underbrace{\begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{bmatrix}}_{\mathbf{x}} = \underbrace{\begin{bmatrix} \mathbf{u}_1 \\ \mathbf{u}_2 \end{bmatrix}}_{\mathbf{u}}, \quad (3.70)$$

Se as matrizes \mathbf{N}_{11} e \mathbf{N}_{22} forem quadradas e inversíveis, a inversa de \mathbf{N} , em bloco, será (STRANG; BORRE, 1997, p. 425):

$$\begin{aligned} \mathbf{N}^{-1} &= \underbrace{\begin{bmatrix} \mathbf{Q}_{\hat{\mathbf{x}}_1} & \mathbf{Q}_{\hat{\mathbf{x}}_1 \hat{\mathbf{x}}_2} \\ \mathbf{Q}_{\hat{\mathbf{x}}_2 \hat{\mathbf{x}}_1} & \mathbf{Q}_{\hat{\mathbf{x}}_2} \end{bmatrix}}_{\mathbf{Q}_{\hat{\mathbf{x}}}} = \\ &= \begin{bmatrix} \mathbf{N}_{11}^{-1} + \mathbf{N}_{11}^{-1} \mathbf{N}_{12} (\mathbf{N}_{22} - \mathbf{N}_{21} \mathbf{N}_{11}^{-1} \mathbf{N}_{12})^{-1} \mathbf{N}_{21} \mathbf{N}_{11}^{-1} & -\mathbf{N}_{11}^{-1} \mathbf{N}_{12} (\mathbf{N}_{22} - \mathbf{N}_{21} \mathbf{N}_{11}^{-1} \mathbf{N}_{12})^{-1} \\ -(\mathbf{N}_{22} - \mathbf{N}_{21} \mathbf{N}_{11}^{-1} \mathbf{N}_{12})^{-1} \mathbf{N}_{21} \mathbf{N}_{11}^{-1} & (\mathbf{N}_{22} - \mathbf{N}_{21} \mathbf{N}_{11}^{-1} \mathbf{N}_{12})^{-1} \end{bmatrix} \end{aligned} \quad (3.71)$$

Usando a expressão (3.71), a solução do sistema (3.69) pode ser explicitada por:

$$\begin{bmatrix} \hat{\mathbf{x}}_1 \\ \hat{\mathbf{x}}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{Q}_{\hat{\mathbf{x}}_1} & \mathbf{Q}_{\hat{\mathbf{x}}_1 \hat{\mathbf{x}}_2} \\ \mathbf{Q}_{\hat{\mathbf{x}}_2 \hat{\mathbf{x}}_1} & \mathbf{Q}_{\hat{\mathbf{x}}_2} \end{bmatrix} \begin{bmatrix} \mathbf{A}_1^T \mathbf{Q}_L^{-1} \mathbf{L} \\ \mathbf{A}_2^T \mathbf{Q}_L^{-1} \mathbf{L} \end{bmatrix}. \quad (3.72)$$

A solução estimada por (3.72) é conhecida na literatura sobre GPS, como solução flutuante (*float solution*). Essa solução, de fato pertence a $\mathbf{R}^{n_1+n_2}$.

A matriz cofatora de $\hat{\mathbf{x}}_2$, isto é, $\mathbf{Q}_{\hat{\mathbf{x}}_2}$ é dada pela expressão:

$$\begin{aligned}\mathbf{Q}_{\hat{\mathbf{x}}_2} &= (\mathbf{N}_{22} - \mathbf{N}_{21}\mathbf{N}_{11}^{-1}\mathbf{N}_{12})^{-1} \\ &= [\mathbf{A}_2^T(\mathbf{Q}_L^{-1} - \mathbf{Q}_L^{-1}\mathbf{A}_1(\mathbf{A}_1^T\mathbf{Q}_L^{-1}\mathbf{A}_1)^{-1}\mathbf{A}_1^T\mathbf{Q}_L^{-1})\mathbf{A}_2]^{-1}.\end{aligned}\quad (3.73)$$

Chamando $\mathbf{Q}^{-1} = \mathbf{Q}_L^{-1} - \mathbf{Q}_L^{-1}\mathbf{A}_1(\mathbf{A}_1^T\mathbf{Q}_L^{-1}\mathbf{A}_1)^{-1}\mathbf{A}_1^T\mathbf{Q}_L^{-1}$, a expressão (3.73) torna-se:

$$\mathbf{Q}_{\hat{\mathbf{x}}_2} = (\mathbf{A}_2^T\mathbf{Q}^{-1}\mathbf{A}_2)^{-1}. \quad (3.74)$$

3.1.5.1 O problema de mínimos quadrados inteiro

Da primeira das equações normais (3.69), tira-se:

$$\hat{\mathbf{x}}_1 = (\mathbf{A}_1^T\mathbf{Q}_L^{-1}\mathbf{A}_1)^{-1}[\mathbf{A}_1^T\mathbf{Q}_L^{-1}\mathbf{L} - \mathbf{A}_1^T\mathbf{Q}_L^{-1}\mathbf{A}_2\mathbf{x}_2]. \quad (3.75)$$

Desenvolvendo a segunda das equações normais de (3.69) e substituindo a equação (3.75), tem-se:

$$\mathbf{A}_2^T(\mathbf{Q}_L^{-1} - \mathbf{Q}_L^{-1}\mathbf{A}_1(\mathbf{A}_1^T\mathbf{Q}_L^{-1}\mathbf{A}_1)^{-1}\mathbf{A}_1^T\mathbf{Q}_L^{-1})\mathbf{A}_2\mathbf{x}_2 = \mathbf{A}_2^T(\mathbf{Q}_L^{-1} - \mathbf{Q}_L^{-1}\mathbf{A}_1(\mathbf{A}_1^T\mathbf{Q}_L^{-1}\mathbf{A}_1)^{-1}\mathbf{A}_1^T\mathbf{Q}_L^{-1})\mathbf{L} \quad (3.76)$$

Note que $(\mathbf{Q}_L^{-1} - \mathbf{Q}_L^{-1}\mathbf{A}_1(\mathbf{A}_1^T\mathbf{Q}_L^{-1}\mathbf{A}_1)^{-1}\mathbf{A}_1^T\mathbf{Q}_L^{-1})$ foi chamado de \mathbf{Q}^{-1} em (3.74). Com essa substituição a equação (3.76), torna-se

$$\mathbf{A}_2^T\mathbf{Q}^{-1}\mathbf{A}_2\mathbf{x}_2 = \mathbf{A}_2^T\mathbf{Q}^{-1}\mathbf{L} \quad (3.77)$$

A solução flutuante explícita para o vetor \mathbf{x}_2 é:

$$\hat{\mathbf{x}}_2 = \mathbf{Q}_2^{-1} \mathbf{u}_2. \quad (3.78)$$

O sistema de equações normais (3.77), corresponde ao sistema de equações normais do problema de mínimos quadrados linear:

$$\min_{\mathbf{x}_2 \in \mathbb{R}^{n_2}} \frac{1}{2} \|\mathbf{W}(\mathbf{A}_2 \mathbf{x}_2 - \mathbf{L})\|_2^2, \quad (3.79)$$

onde:

$$\mathbf{W}^T \mathbf{W} = \mathbf{Q}^{-1}.$$

Impondo a condição $\mathbf{x}_2 \in \mathbb{Z}^{n_2}$, o problema torna-se:

$$\min_{\mathbf{x}_2 \in \mathbb{Z}^{n_2}} \frac{1}{2} \|\mathbf{W}(\mathbf{A}_2 \mathbf{x}_2 - \mathbf{L})\|_2^2, \quad (3.80)$$

e passa a ser conhecido por **problema de mínimos quadrados inteiro**.

Usando a solução flutuante (3.78), o problema (3.80) pode ser colocado na forma equivalente (fig. 3.4):

$$\min_{\mathbf{x}_2 \in \mathbb{Z}^{n_2}} \frac{1}{2} \|\mathbf{W} \mathbf{A}_2 (\mathbf{x}_2 - \hat{\mathbf{x}}_2)\|_2^2. \quad (3.81)$$

Freqüentemente o problema de mínimos quadrados inteiro (3.81), aparece na literatura com a forma:

$$\min_{\mathbf{a} \in \mathbb{Z}^{n_2}} \chi^2(\mathbf{a}) = (\mathbf{a} - \hat{\mathbf{a}})^T \mathbf{Q}_2^{-1} (\mathbf{a} - \hat{\mathbf{a}}), \quad (3.82)$$

onde:

$$\hat{\mathbf{a}} = \hat{\mathbf{x}}_2; \quad \mathbf{a} = \mathbf{x}_2;$$

$$\mathbf{Q}_2^{-1} = \mathbf{A}_2^T \mathbf{Q}^{-1} \mathbf{A}_2 \quad (\text{ver equação 3.74}),$$

com:

$$\mathbf{Q}^{-1} = \mathbf{W}^T \mathbf{W} = \mathbf{Q}_L^{-1} - \mathbf{Q}_L^{-1} \mathbf{A}_1 (\mathbf{A}_1^T \mathbf{Q}_L^{-1} \mathbf{A}_1)^{-1} \mathbf{A}_1^T \mathbf{Q}_L^{-1}.$$

Obs.: a notação (3.82) será algumas vezes usada no cap.4. A constante $(\frac{1}{2})$ introduzido anteriormente, apenas por conveniência, foi retirada em (3.82).

Com a expressão (3.83), fica ilustrado que os elementos do vetor de inteiros estimado ($\tilde{\mathbf{a}}$), são obtidos dos inteiros mais próximos de cada elemento de $\hat{\mathbf{a}}$.

Se a matriz \mathbf{Q}_a^{-1} não for diagonal, o problema (3.82) torna-se bastante difícil, e não existe um método de resolução direta do problema como é o caso do problema de mínimos quadrados linear. Os métodos existentes, de uma forma ou de outra, realizam buscas dos vetores inteiros (\mathbf{a}), dentro de um determinado espaço, até encontrar o vetor inteiro ($\tilde{\mathbf{a}}$) minimizador da forma quadrática (3.82).

Uma vez encontrado o vetor $\tilde{\mathbf{a}} = \tilde{\mathbf{x}}_2$, o vetor real correspondente ($\tilde{\mathbf{x}}_1$), é estimado resolvendo-se o seguinte sistema (ver equação 3.75):

$$\mathbf{A}_1^T \mathbf{Q}_L^{-1} \mathbf{A}_1 \tilde{\mathbf{x}}_1 = \mathbf{A}_1^T \mathbf{Q}_L^{-1} \mathbf{L} - \mathbf{A}_1^T \mathbf{Q}_L^{-1} \tilde{\mathbf{x}}_2. \quad (3.84)$$

3.1.6 Problema de Mínimos Quadrados Linear com Injunções de Igualdade

No ajustamento de observações geodésicas é muito comum situações onde se adota valores conhecidos para alguns parâmetros (exemplo: coordenadas de estações, ambigüidades, etc.), ou para uma grandeza relacionada aos parâmetros através de uma função (exemplo, condição de ângulo, de azimuth, ou de distância).

As injunções¹² aos parâmetros podem ser representadas por uma função não-linear implícita, como:

$$\mathbf{h}(\mathbf{x}) = \mathbf{0}; \quad (3.85)$$

onde:

$$\mathbf{h} : \mathbf{R}^n \rightarrow \mathbf{R}^t, \quad t \leq n.$$

¹² No Brasil o termo em inglês *constraint* é traduzido em Geodésia como injunção. Em Matemática, o termo é conhecido como restrição, e há áreas onde o termo é conhecido como vínculo.

O ajustamento por mínimos quadrados com as injunções (3.85) poderia ser realizado por métodos de minimização sem restrição, utilizando a função Lagrangeano.

Sendo \mathbf{x}_0 um ponto suficientemente próximo da solução do problema de mínimos quadrados não-linear, com as injunções (3.85), o problema pode ser trocado pelo problema de mínimos quadrados linear, com injunção linear (obtido por meio da linearização no ponto \mathbf{x}_0), dado por:

$$\min_{\mathbf{x} \in \mathbb{R}^n} \frac{1}{2} \|\mathbf{W}(\mathbf{Ax} - \mathbf{L})\|_2^2 \quad (3.86)$$

$$\text{sujeito a: } \mathbf{Gx} = \mathbf{a},$$

onde:

$\mathbf{a} \in \mathbb{R}^t$: vetor com elementos conhecidos (injunções);

$\mathbf{G} \in \mathbb{R}^{t \times n}$, com $t \leq n$: matriz Jacobiano de \mathbf{h} em \mathbf{x}_0 .

O Lagrangeano do problema (3.86) pode ser definido como:

$$f(\mathbf{x}, \mathbf{k}) = \frac{1}{2} \mathbf{x}^T \mathbf{A}^T \mathbf{Q}_L^{-1} \mathbf{Ax} - \mathbf{x}^T \mathbf{A}^T \mathbf{Q}_L^{-1} \mathbf{L} + \frac{1}{2} \mathbf{L}^T \mathbf{Q}_L^{-1} \mathbf{L} - (\mathbf{Gx} - \mathbf{a})^T \mathbf{k}, \quad (3.87)$$

onde:

$\mathbf{k} \in \mathbb{R}^t$: é o vetor dos multiplicadores de Lagrange.

Derivando $f(\mathbf{x}, \mathbf{k})$ em relação ao vetor \mathbf{x} , em relação ao vetor \mathbf{k} , e igualando a zero (condições de otimalidade de primeira ordem), obtém-se os sistemas:

$$\mathbf{A}^T \mathbf{Q}_L^{-1} \mathbf{Ax} - \mathbf{A}^T \mathbf{Q}_L^{-1} \mathbf{L} - \mathbf{G}^T \mathbf{k} = \mathbf{0}, \quad (3.88)$$

$$\mathbf{Gx} - \mathbf{a} = \mathbf{0}. \quad (3.89)$$

Se a matriz \mathbf{A} tem posto completo, então:

$$\mathbf{x} = (\mathbf{A}^T \mathbf{Q}_L^{-1} \mathbf{A})^{-1} \mathbf{A}^T \mathbf{Q}_L^{-1} \mathbf{L} + (\mathbf{A}^T \mathbf{Q}_L^{-1} \mathbf{A})^{-1} \mathbf{G}^T \mathbf{k}. \quad (3.90)$$

O primeiro termo do lado direito corresponde à solução sem a influência das injunções, e o segundo corresponde exatamente à influência das injunções. Assim,

$$\mathbf{x} = \hat{\mathbf{x}} + \Delta \mathbf{x}. \quad (3.91)$$

Levando (3.90) em (3.89), tem-se:

$$\mathbf{G}(\mathbf{A}^T \mathbf{Q}_L^{-1} \mathbf{A})^{-1} \mathbf{A}^T \mathbf{Q}_L^{-1} \mathbf{L} + \mathbf{G}(\mathbf{A}^T \mathbf{Q}_L^{-1} \mathbf{A})^{-1} \mathbf{G}^T \mathbf{k} = \mathbf{a} \quad (3.92)$$

Se posto(\mathbf{G}) = t , então:

$$\mathbf{k} = [\mathbf{G}(\mathbf{A}^T \mathbf{Q}_L^{-1} \mathbf{A})^{-1} \mathbf{G}^T]^{-1} (\mathbf{a} - \mathbf{G}(\mathbf{A}^T \mathbf{Q}_L^{-1} \mathbf{A})^{-1} \mathbf{A}^T \mathbf{Q}_L^{-1} \mathbf{L}), \quad (3.93)$$

e a influência das injunções é calculada por:

$$\Delta \mathbf{x} = (\mathbf{A}^T \mathbf{Q}_L^{-1} \mathbf{A})^{-1} \mathbf{G}^T [\mathbf{G}(\mathbf{A}^T \mathbf{Q}_L^{-1} \mathbf{A})^{-1} \mathbf{G}^T]^{-1} (\mathbf{a} - \mathbf{G}(\mathbf{A}^T \mathbf{Q}_L^{-1} \mathbf{A})^{-1} \mathbf{A}^T \mathbf{Q}_L^{-1} \mathbf{L}). \quad (3.94)$$

As expressões (3.93) e (3.94) são derivadas para uma matriz $\mathbf{G} \in \mathbb{R}^{t \times n}$, com posto(\mathbf{G}) = t , qualquer. Nesse trabalho é de interesse o caso particular $\mathbf{G} = [\mathbf{0} \ \mathbf{I}]$ por causa da natureza inteira das ambigüidades.

Particionando-se o vetor \mathbf{x} em seu componente real (\mathbf{x}_1) e seu componente inteiro (\mathbf{x}_2), como foi na equação (3.67), e também a matriz de coeficientes dos parâmetros injuncionados $\mathbf{G} = [\mathbf{G}_1 \ \mathbf{G}_2]$, onde \mathbf{G}_1 tem os coeficientes das injunções de \mathbf{x}_1 e \mathbf{G}_2 os coeficientes das injunções de \mathbf{x}_2 .

Com os particionamentos as equações (3.88) e (3.89) tornam-se, respectivamente:

$$\begin{bmatrix} \mathbf{A}_1^T \mathbf{Q}_L^{-1} \mathbf{A}_1 & \mathbf{A}_1^T \mathbf{Q}_L^{-1} \mathbf{A}_2 \\ \mathbf{A}_2^T \mathbf{Q}_L^{-1} \mathbf{A}_1 & \mathbf{A}_2^T \mathbf{Q}_L^{-1} \mathbf{A}_2 \end{bmatrix} \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{A}_1^T \mathbf{Q}_L^{-1} \mathbf{L} \\ \mathbf{A}_2^T \mathbf{Q}_L^{-1} \mathbf{L} \end{bmatrix} + \begin{bmatrix} \mathbf{G}_1^T \mathbf{k} \\ \mathbf{G}_2^T \mathbf{k} \end{bmatrix}, \quad (3.95)$$

$$[\mathbf{G}_1 \ \mathbf{G}_2] \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{bmatrix} = \mathbf{a}. \quad (3.96)$$

O vetor \mathbf{a} (vetor de ambigüidades conhecidas), representa as injunções ao vetor \mathbf{x}_2 , logo $\mathbf{G}_1 = \mathbf{0}$ e $\mathbf{G}_2 = \mathbf{I}$.

Lembrando da solução flutuante (3.72), e considerando \mathbf{G}_1 e \mathbf{G}_2 conforme acima, o sistema (3.95) pode ser representado como:

$$\left. \begin{aligned} \mathbf{x}_1 &= \hat{\mathbf{x}}_1 + \mathbf{Q}_{\hat{\mathbf{x}}_1 \hat{\mathbf{x}}_2} \mathbf{k} \\ \mathbf{x}_2 &= \hat{\mathbf{x}}_2 + \mathbf{Q}_{\hat{\mathbf{x}}_2} \mathbf{k} \end{aligned} \right\} \quad (3.97)$$

Como $\mathbf{Q}_{\hat{\mathbf{x}}_2}$ é inversível (ver equação 3.71), então:

$$\mathbf{k} = \mathbf{Q}_{\hat{\mathbf{x}}_2}^{-1} (\mathbf{x}_2 - \hat{\mathbf{x}}_2). \quad (3.98)$$

Levando essa última na primeira equação de (3.97) e substituindo $\mathbf{a} = \mathbf{x}_2$, resulta:

$$\tilde{\mathbf{x}}_1 = \hat{\mathbf{x}}_1 - \mathbf{Q}_{\hat{\mathbf{x}}_1 \hat{\mathbf{x}}_2} \mathbf{Q}_{\hat{\mathbf{x}}_2}^{-1} (\hat{\mathbf{x}}_2 - \mathbf{a}). \quad (3.99)$$

A equação (3.99), mostra a influência das injunções \mathbf{a} , sobre a solução $\hat{\mathbf{x}}_1$. Para aplicá-la não foi imposta a condição de que o vetor \mathbf{a} tem que ser inteiro, basta que as matrizes sejam $\mathbf{G}_1 = \mathbf{0}$ e $\mathbf{G}_2 = \mathbf{I}$. Para os propósitos desse trabalho, o vetor \mathbf{a} interpreta as ambigüidades GPS, o vetor $\hat{\mathbf{x}}_2$ interpreta as ambigüidades reais (solução flutuante) e o vetor $\hat{\mathbf{x}}_1$ a solução flutuante das coordenadas da estação incógnita da linha de base.

A equação (3.99) é útil para a realização de algumas análises, servindo também como um caminho alternativo para o cálculo de $\tilde{\mathbf{x}}_1$ (ver equação 3.84).

3.1.7 Problema de Mínimos Quadrados Não-Linear

O modelo das equações de observação que vem sendo tratado ao longo desse capítulo é (ver equação 3.1):

$$\mathbf{L}_a = \mathbf{F}(\mathbf{x}_a) . \quad (3.100)$$

Substituindo o vetor das observações ajustadas (\mathbf{L}_a), pelo vetor das observações (\mathbf{L}_b), o modelo (3.100) pode ser colocado na forma (ver equação 1.3):

$$\mathbf{r}(\mathbf{x}) = \mathbf{F}(\mathbf{x}) - \mathbf{L}_b , \quad (3.101)$$

onde :

$\mathbf{r} : \mathbf{R}^n \rightarrow \mathbf{R}^m$: função residual;

$\mathbf{F} : \mathbf{R}^n \rightarrow \mathbf{R}^m$: modelo funcional das observações;

$\mathbf{L}_b \in \mathbf{R}^m$: vetor das observações;

$\mathbf{x} \in \mathbf{R}^n$: vetor dos parâmetros.

O problema de mínimos quadrados para (3.101) é definido como:

$$\min_{\mathbf{x} \in \mathbf{R}^n} f(\mathbf{x}) = \frac{1}{2} \|\mathbf{r}(\mathbf{x})\|_{\mathbf{Q}_{L_b}^{-1}}^2 . \quad (3.102)$$

Se $\mathbf{r}(\mathbf{x})$ é linear, o modelo (3.101) pode ser representado por $\mathbf{V} = \mathbf{A}\mathbf{x} - \mathbf{L}$, e o problema passa a ser o de mínimos quadrados linear, já discutido nas seções anteriores. Sendo $\mathbf{r}(\mathbf{x})$ não-linear, o problema (3.102), é um problema de mínimos quadrados não-linear. A solução para esse caso pode ser obtida por meio dos métodos apresentados a seguir.

3.1.7.1 Método de Newton

O método de Newton é um dos métodos mais clássicos para resolver

problemas de minimização irrestrita (sem injunções). A cada iteração ele utiliza uma aproximação quadrática da função $f(\mathbf{x})$ em um ponto \mathbf{x}_k e obtém o minimizador desta quadrática.

A aproximação quadrática de (3.102) no ponto \mathbf{x}_k , pode ser representada por:

$$f_c(\mathbf{x}) = f(\mathbf{x}_k) + \nabla f(\mathbf{x}_k)(\mathbf{x} - \mathbf{x}_k) + \frac{1}{2} (\mathbf{x} - \mathbf{x}_k)^T \nabla^2 f(\mathbf{x}_k)(\mathbf{x} - \mathbf{x}_k) \quad (3.103)$$

onde:

$$f(\mathbf{x}_k) = \frac{1}{2} \mathbf{r}(\mathbf{x}_k)^T \mathbf{Q}_{L_b}^{-1} \mathbf{r}(\mathbf{x}_k) : \text{valor da função objetivo no ponto } \mathbf{x}_k;$$

$$\nabla f(\mathbf{x}_k) = \mathbf{J}(\mathbf{x}_k)^T \mathbf{Q}_{L_b}^{-1} \mathbf{r}(\mathbf{x}_k) : \text{gradiente de } f(\mathbf{x}_k);$$

$$\nabla^2 f(\mathbf{x}_k) = \mathbf{J}(\mathbf{x}_k)^T \mathbf{Q}_{L_b}^{-1} \mathbf{J}(\mathbf{x}_k) + \mathbf{S}(\mathbf{x}_k) : \text{Hessiana de } f(\mathbf{x}_k), \text{ com } \mathbf{S}(\mathbf{x}_k) \text{ contendo as} \\ \text{informações de segunda ordem;}$$

$$\mathbf{J}(\mathbf{x}_k) : \text{é a matriz Jacobiano de } \mathbf{r}(\mathbf{x}) \text{ no ponto } \mathbf{x}_k.$$

O gradiente da quadrática (3.103) é:

$$\nabla f_c(\mathbf{x}) = \nabla f(\mathbf{x}_k) + \nabla^2 f(\mathbf{x}_k)(\mathbf{x} - \mathbf{x}_k), \quad (3.104)$$

e da condição de otimalidade $\nabla f_c(\mathbf{x}) = 0$, com a substituição de $\nabla f(\mathbf{x}_k)$ e $\nabla^2 f(\mathbf{x}_k)$ na equação acima, tem-se:

$$[\mathbf{J}(\mathbf{x}_k)^T \mathbf{Q}_{L_b}^{-1} \mathbf{J}(\mathbf{x}_k) + \mathbf{S}(\mathbf{x}_k)](\mathbf{x}_{k+1} - \mathbf{x}_k) = - \mathbf{J}(\mathbf{x}_k)^T \mathbf{Q}_{L_b}^{-1} \mathbf{r}(\mathbf{x}_k) \quad (3.105)$$

O vetor $\mathbf{d}_k = \mathbf{x}_{k+1} - \mathbf{x}_k$, solução do sistema (3.105), é o passo de Newton do problema de mínimos quadrados não-linear (3.102). Se a matriz $[\mathbf{J}(\mathbf{x}_k)^T \mathbf{Q}_{L_b}^{-1} \mathbf{J}(\mathbf{x}_k) + \mathbf{S}(\mathbf{x}_k)]$ é definida positiva, então o passo de Newton é uma direção de descida¹³ de $f(\mathbf{x})$ em \mathbf{x}_k , pois:

¹³ Sobre direções de descida e condições de otimalidade de primeira e segunda ordem dos problemas de minimização, consultar, por exemplo, MARTINEZ e SANTOS (1995, p. 100, p.12), FRIEDLANDER (1994, p. 29, p. 17).

$$\begin{aligned}
 -\mathbf{d}_k^T [\mathbf{J}(\mathbf{x}_k)^T \mathbf{Q}_{L_b}^{-1} \mathbf{J}(\mathbf{x}_k) + \mathbf{S}(\mathbf{x}_k)] \mathbf{d}_k &= \mathbf{d}_k^T \mathbf{J}(\mathbf{x}_k)^T \mathbf{Q}_{L_b}^{-1} \mathbf{r}(\mathbf{x}_k) \\
 &= \nabla f(\mathbf{x}_k)^T \mathbf{d}_k < 0
 \end{aligned}
 \tag{3.106}$$

Quando a matriz Hessiana é definida positiva na vizinhança da solução \mathbf{x}_a (\mathbf{x}_a é o vetor dos parâmetros ajustados), e \mathbf{x}_0 é tomado suficientemente próximo da solução, o método de Newton converge quadraticamente (prova: ver LUENBERG, 1973, p. 156).

3.1.7.2 Método de Gauss-Newton

Em um ponto aproximado (\mathbf{x}_k), a função residual não-linear ($\mathbf{r}(\mathbf{x})$), pode ser aproximada pela seguinte função afim:

$$\mathbf{V}(\mathbf{x}_k) = \mathbf{F}(\mathbf{x}_k) - \mathbf{L}_b + \frac{\partial \mathbf{r}}{\partial \mathbf{x}} \Big|_{\mathbf{x}_k} (\mathbf{x} - \mathbf{x}_k) .
 \tag{3.107}$$

Note que $\mathbf{F}(\mathbf{x}_k)$ representa o vetor de observações calculadas no ponto \mathbf{x}_k e $\frac{\partial \mathbf{r}}{\partial \mathbf{x}} = \frac{\partial \mathbf{F}}{\partial \mathbf{x}} = \mathbf{J}(\mathbf{x}_k)$ é a matriz Jacobiano de $\mathbf{r}(\mathbf{x})$ calculada no ponto \mathbf{x}_k . A equação (3.107) pode ser reescrita na forma (ver seção 3.1.1):

$$\mathbf{V} = \mathbf{A} \Delta \mathbf{x} - \mathbf{L},
 \tag{3.108}$$

onde:

$$\mathbf{A} = \mathbf{J}(\mathbf{x}_k) ; \quad \mathbf{L} = \mathbf{L}_b - \mathbf{F}(\mathbf{x}_k); \quad \text{e} \quad \Delta \mathbf{x} = \mathbf{x} - \mathbf{x}_k = \mathbf{d}_k.$$

O vetor $\Delta \mathbf{x}$ do método de Gauss-Newton, é obtido da solução do problema de mínimos quadrados linear

$$\min_{\Delta \mathbf{x} \in \mathcal{R}^p} \frac{1}{2} \| \mathbf{A} \Delta \mathbf{x} - \mathbf{L} \|^2_{\mathbf{Q}_{L_b}^{-1}} .
 \tag{3.109}$$

Uma vez encontrado $\Delta \hat{\mathbf{x}}$ solução do problema (3.109), \mathbf{x}_{k+1} é calculado por:

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \Delta \hat{\mathbf{x}}. \quad (3.110)$$

As equações (3.107) a (3.110) caracterizam o método de Gauss-Newton.

Algoritmo 3.7 (Gauss-Newton). Dada a função residual $\mathbf{r} : \mathbb{R}^n \rightarrow \mathbb{R}^m$ conforme definida em (3.101), um ponto aproximado $\mathbf{x}_k \in \mathbb{R}^n$, e a matriz dos pesos $\mathbf{Q}_{L_b}^{-1} \in \mathbb{R}^{m \times m}$ simétrica e definida positiva, em cada iteração do método, \mathbf{x}_{k+1} é obtido por:

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \arg \min_{\Delta \mathbf{x} \in \mathbb{R}^n} \frac{1}{2} \|\mathbf{J}(\mathbf{x}_k) \Delta \mathbf{x} - (\mathbf{L}_b - \mathbf{F}(\mathbf{x}_k))\|_{\mathbf{Q}_{L_b}}^2 \quad (3.111)$$

até que

$$\|\mathbf{x}_{k+1} - \mathbf{x}_k\| < tol_{\text{passo}}; \text{ ou} \quad (3.112)$$

$$\|\nabla f(\mathbf{x}_{k+1})\| = \|\mathbf{J}(\mathbf{x}_{k+1})^T \mathbf{Q}_{L_b}^{-1} \mathbf{r}(\mathbf{x}_{k+1})\| < tol_{\text{grad}}; \text{ ou} \quad (3.113)$$

$$k > tol_{\text{iter}}. \quad (3.114)$$

A condição (3.114) limita o número de iterações, a qual pode indicar a divergência do método, ou indicar que o tamanho do passo $\|\mathbf{x}_{k+1} - \mathbf{x}_k\|$ não está atingindo o mínimo exigido pela condição tol_{passo} .

A equação (3.111) pode ser colocada em termos do sistema de equações normais:

$$[\mathbf{J}(\mathbf{x}_k)^T \mathbf{Q}_{L_b}^{-1} \mathbf{J}(\mathbf{x}_k)](\mathbf{x}_{k+1} - \mathbf{x}_k) = -\mathbf{J}(\mathbf{x}_k)^T \mathbf{Q}_{L_b}^{-1} \mathbf{r}(\mathbf{x}_k). \quad (3.115)$$

Comparando a (3.115) com a (3.105), nota-se que a única diferença entre o passo do método de Gauss-Newton e o de Newton, é pela influência da matriz $\mathbf{S}(\mathbf{x}_k)$ de informações de segunda ordem da Hessiana.

Quando $\mathbf{S}(\mathbf{x}_a) = \mathbf{0}$ (ver Hessiana da equação (3.103)) o método de

Gauss-Newton, sob certas condições, é quadraticamente convergente. Se $S(\mathbf{x}_a)$ for relativamente pequena quando comparada a $[\mathbf{J}(\mathbf{x}_{k+1})^T \mathbf{Q}_{L_b}^{-1} \mathbf{J}(\mathbf{x}_{k+1})]$, o método de Gauss-Newton é linearmente convergente e, se $S(\mathbf{x}_a)$ for muito grande, o método pode não convergir. Detalhes sobre as informações de convergência do método de Gauss-Newton são demonstrados em DENNIS JR e SCHNABEL (1983, p.222).

3.1.7.3 Método de Levenberg-Marquardt

Pode existir situações onde a matriz Jacobiano $\mathbf{J}(\mathbf{x}_k)$ tenha deficiência de posto. Neste caso, a matriz $[\mathbf{J}(\mathbf{x}_k)^T \mathbf{Q}_{L_b}^{-1} \mathbf{J}(\mathbf{x}_k)]$ será singular e o passo de método de Gauss-Newton não estará bem definido.

Uma estratégia proposta por LEVENBERG (1944)¹⁴ e depois por MARQUARDT (1963)¹⁵, citado por DENNIS JR e SCHANABEL (1983, p.227), é usar:

$$[\mathbf{J}(\mathbf{x}_k)^T \mathbf{Q}_{L_b}^{-1} \mathbf{J}(\mathbf{x}_k) + \mu_k \mathbf{I}](\mathbf{x}_{k+1} - \mathbf{x}_k) = - \mathbf{J}(\mathbf{x}_k)^T \mathbf{Q}_{L_b}^{-1} \mathbf{r}(\mathbf{x}_k), \quad (3.116)$$

onde: μ_k é um escalar positivo.

Se $\mu_k = 0$, o passo $\mathbf{d}_k = (\mathbf{x}_{k+1} - \mathbf{x}_k)$ é o passo do método de Gauss-Newton e se $\mu_k \rightarrow \infty$, $\|\mathbf{d}_k\| \rightarrow 0$, e a direção do passo tende para $-\nabla f(\mathbf{x}_k)$.

O método de Levenberg-Marquardt é um método baseado em região de confiança. Isto porque, (3.116) é a solução de um problema com a injunção de que o comprimento do passo $\|\mathbf{d}_k\|$, tem que ser inferior a um escalar (δ) relacionado a μ_k . O escalar (δ) define a região de confiança. Ver GILL; MURRAY e WRIGHT (1981, p. 136), DENNIS JR e SCHANABEL (1983, p.227) e MARTINEZ e SANTOS (1995, p. 126).

O problema de mínimos quadrados com região de confiança é definido como:

¹⁴ LEVENBERG, K. "A method for the solution of certain problems in least squares," *Quart. Appl. Math.* 2, 1944, p. 164 -168.

¹⁵ MARQUARDT, D. "An algorithm for least - squares estimation of nonlinear parameters," *SIAM J. Appl. Math.* 11, 1963, p. 431 - 441.

$$\min_{\mathbf{d}_k \in \mathbb{R}^n} \frac{1}{2} \|\mathbf{J}(\mathbf{x}_k)\mathbf{d}_k + \mathbf{r}(\mathbf{x}_k)\|_{\mathbf{Q}_{L_b}^{-1}}^2, \quad (3.117)$$

sujeito a : $\|\mathbf{d}_k\| < \delta$.

Existem várias estratégias para implementar μ_k no algoritmo de Levenberg-Marquardt. Veja, por exemplo, em: PRESS et al., (1986,p.254), DENNIS JR e SCHNABEL (1983, seção 6.4) e GARNÉS (1996, p. 137).

A razão de convergência do método de Levenberg-Marquardt é semelhante a do método de Gauss-Newton.

Experiências realizadas por GARNÉS (1996, p.139), GARNÉS; SAMPAIO e DALMOLIN (1997, p.32) em problemas práticos de Geodésia, mostraram que o método se comporta muito bem, inclusive com características de convergência Global, onde o método de Gauss-Newton falhou, o método de Levenberg-Marquardt obteve sucesso.

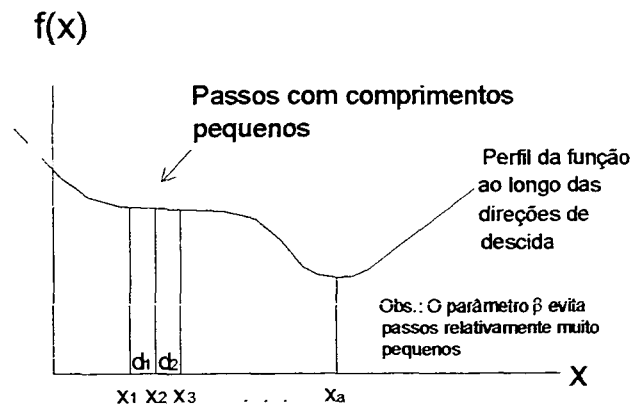
3.1.7.4 Método de Newton globalizado

O passo do método de Newton quando aplicado ao problema de mínimos quadrados não-linear, equação (3.105), poderá não estar bem definido se a matriz Hessiana ($\nabla^2 f(\mathbf{x}_k)$) não for definida positiva. Uma estratégia semelhante ao método de Levenberg-Marquardt poderia ser aplicada a fim de deixar a matriz Hessiana definida positiva em \mathbf{x}_k . No entanto, para garantir a convergência global¹⁶, é necessário o cumprimento de 3 condições ilustradas na figura 3.5 a seguir.

¹⁶ Um método é dito ter convergência global, quando independente do ponto inicial \mathbf{x}_k , ele converge para um ponto estacionário, isto é, $\nabla f(\mathbf{x}_*) = \mathbf{0}$ (\mathbf{x}_* representa o ponto estacionário).

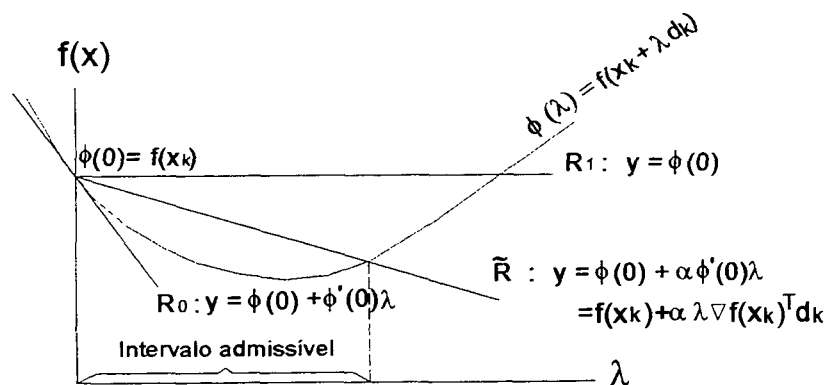
FIGURA 3.5 – CONDIÇÕES PARA CONVERGÊNCIA GLOBAL

3.5.(a) – Condição do tamanho do passo



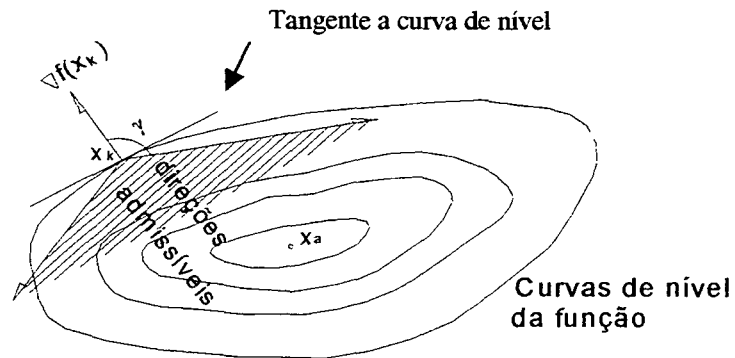
$$\|d_k\| \geq \beta \|\nabla f(x_k)\|, \quad \beta > 0$$

3.5.(b) – Condição de Armijo



$$f(x_k + \lambda d_k) < f(x_k) + \alpha \lambda \nabla f(x_k)^T d_k$$

3.5.(c) – Condição de Ângulo



$$\nabla f(\mathbf{x}_k)^T \mathbf{d}_k \leq -\theta \|\nabla f(\mathbf{x}_k)\| \|\mathbf{d}_k\|,$$

$$\text{onde: } \cos \gamma = -\theta$$

A condição do tamanho do passo (fig.3.5a), $\|\mathbf{d}_k\| \geq \beta \|\nabla f(\mathbf{x}_k)\|$ com $\beta > 0$ escolhido adequadamente, evita passos muito pequenos em relação ao tamanho do gradiente.

A condição de Armijo, $f(\mathbf{x}_{k+1}) < f(\mathbf{x}_k) + \alpha \lambda_k \nabla f(\mathbf{x}_k)^T \mathbf{d}_k$, impõe que λ_k esteja no intervalo admissível (fig. 3.5b) e com isso assegura um novo \mathbf{x}_{k+1} , com um decréscimo na função abaixo da reta \tilde{R} . Essa condição é também conhecida como condição de decréscimo suficiente.

A condição de ângulo, $\nabla f(\mathbf{x}_k)^T \mathbf{d}_k \leq -\theta \|\nabla f(\mathbf{x}_k)\| \|\mathbf{d}_k\|$, impõe que o passo seja dado num ângulo maior que γ (com $\gamma > 90^\circ$ em relação a $\nabla f(\mathbf{x}_k)$) (figura 3.5c). Isso impede o método de andar aproximadamente sobre a mesma curva de nível.

Com essas três condições o algoritmo fica bem definido e para algum valor de k , se a sequência $\{\mathbf{x}_k\}$ é limitada, o limite é um ponto estacionário, ou seja, $\nabla f(\mathbf{x}_k) = \mathbf{0}$.

O algoritmo a seguir, sugerido por MARTINEZ e SANTOS (1995, p.104), mostra como aplicar essas 3 condições.

Algoritmo 3.8 (Descida com condições de globalização). Sejam $\mathbf{x}_k \in \mathbb{R}^n$, $\alpha \in (0,1)$, $\beta > 0$, $\theta \in (0,1)$. Os passos para determinar \mathbf{x}_{k+1} são:

1. Se $\nabla f(\mathbf{x}_k) = 0$, parar.
2. Escolher $\mathbf{d}_k \in \mathbb{R}^n$ tal que

$$\|\mathbf{d}_k\| \geq \beta \|\nabla f(\mathbf{x}_k)\|$$

$$\nabla f(\mathbf{x}_k)^T \mathbf{d}_k \leq -\theta \|\nabla f(\mathbf{x}_k)\| \|\mathbf{d}_k\|$$
3. $\lambda = 1$.
4. Enquanto $f(\mathbf{x}_k + \lambda \mathbf{d}_k) > f(\mathbf{x}_k) + \alpha \lambda \nabla f(\mathbf{x}_k)^T \mathbf{d}_k$
 escolher novo $\lambda \in [0,1\lambda, 0,9\lambda]$
5. $\mathbf{x}_{k+1} = \mathbf{x}_k + \lambda \mathbf{d}_k$.

Conforme mencionado anteriormente, se a matriz Hessiana não for definida positiva, pode-se usar a estratégia de aumentar os valores da diagonal principal, até torná-la definida positiva (e conseqüentemente diagonalmente dominante). Com a observância de que a estratégia tem o preço de tornar o algoritmo mais lento por causa da degradação da direção de descida. Assim:

$$\mathbf{B}_k = \nabla^2 f(\mathbf{x}_k) + \mu_k \mathbf{I}, \quad \text{com } \mu_k > 0. \quad (3.118)$$

Substituindo a Hessiana $\nabla^2 f(\mathbf{x}_k) = [\mathbf{J}(\mathbf{x}_k)^T \mathbf{Q}_{L_b}^{-1} \mathbf{J}(\mathbf{x}_k) + \mathbf{S}(\mathbf{x}_k)]$ por \mathbf{B}_k em (3.105), tem-se

$$\mathbf{B}_k \mathbf{d}_k = -\mathbf{J}(\mathbf{x}_k)^T \mathbf{Q}_{L_b}^{-1} \mathbf{r}(\mathbf{x}_k). \quad (3.119)$$

Por ser a decomposição de Cholesky a forma mais barata para verificar se \mathbf{B}_k é definida positiva (ver seção 3.1.3.1), a idéia é tentá-la e, se não for possível, aumentar o valor de μ_k até obter sucesso. Com isso, a solução \mathbf{d}_k de (3.119) será uma direção de descida.

MARTINEZ e SANTOS (1995, p.108), apresentam o algoritmo a seguir com essas idéias incorporadas, recebendo a denominação de método de Newton com busca linear. Nesse trabalho, no entanto, será chamado de método de Newton modificado globalizado.

Algoritmo 3.9 (Newton modificado globalizado). Dados $\alpha \in (0,1)$, $\beta > 0$, $\theta \in (0,1)$ e $\mathbf{x}_k \in R^n$

1. Se $\nabla f(\mathbf{x}_k) = \mathbf{0}$, parar.
2. Tentar a decomposição de Cholesky $\nabla^2 f(\mathbf{x}_k) = \mathbf{L} \mathbf{L}^T$
3. Se houve sucesso em (2), obter \mathbf{d}_k resolvendo

$$\mathbf{L} \mathbf{y} = -\nabla f(\mathbf{x}_k)$$

$$\mathbf{L}^T \mathbf{d}_k = \mathbf{y}.$$

4. Se (2) fracassou, definir $\mathbf{B}_k = \nabla^2 f(\mathbf{x}_k) + \mu \mathbf{I}$, $\mu > 0$, de maneira que \mathbf{B}_k seja definida positiva. Realizar a decomposição de Cholesky $\mathbf{B}_k = \mathbf{L} \mathbf{L}^T$ e resolver os sistemas

$$\mathbf{L} \mathbf{y} = -\nabla f(\mathbf{x}_k)$$

$$\mathbf{L}^T \mathbf{d}_k = \mathbf{y}.$$

5. Se $\nabla f(\mathbf{x}_k)^T \mathbf{d}_k > -\theta \|\nabla f(\mathbf{x}_k)\| \|\mathbf{d}_k\|$, fazer $\mu \leftarrow \max \{2\mu, 10\}$ e repetir o passo (4) como se tivesse havido fracasso na decomposição de Cholesky.
6. Se $\|\mathbf{d}_k\| < \beta \|\nabla f(\mathbf{x}_k)\|$, corrigir:

$$\mathbf{d}_k \leftarrow \beta \frac{\mathbf{d}_k}{\|\mathbf{d}_k\|} \|\nabla f(\mathbf{x}_k)\|$$

7. Obter λ de modo a satisfazer a condição de Armijo

$$f(\mathbf{x}_k + \lambda \mathbf{d}_k) \leq f(\mathbf{x}_k) + \alpha \lambda \nabla f(\mathbf{x}_k)^T \mathbf{d}_k,$$

definir

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \lambda \mathbf{d}_k$$

e voltar para (1).

Os parâmetros α , β e θ utilizados no algoritmo 3.9, devem ter valores de maneira a permitir a boa performance do método. DENNIS JR e SCHNABEL (1983,

p.126) utilizam para α , $\alpha = 1.10^{-4}$.

Os parâmetros θ e β podem ser definidos de modo a permitir que o método de Newton seja usado naturalmente próximo da solução, aproveitando assim, a sua razão de convergência quadrática. Os limites naturais para θ e β para um passo do método de Newton, são definidos como exposto na sequência.

O passo natural de Newton é definido pela solução do sistema

$$\nabla^2 f(\mathbf{x}_k) \mathbf{d}_k = - \nabla f(\mathbf{x}_k) . \quad (3.120)$$

Usando as propriedades de uma norma compatível, tem-se:

$$\|\mathbf{d}_k\| \geq \frac{1}{\|\nabla^2 f(\mathbf{x}_k)\|} \|\nabla f(\mathbf{x}_k)\| , \text{ logo } \beta = \frac{1}{\|\nabla^2 f(\mathbf{x}_k)\|} . \quad (3.121)$$

Com a norma-2, β é o inverso do maior autovalor da matriz Hessiana.

Pré-multiplicando (3.120) pelo inverso da matriz Hessiana

$$[\nabla^2 f(\mathbf{x}_k)]^{-1} \nabla^2 f(\mathbf{x}_k) \mathbf{d}_k = - [\nabla f(\mathbf{x}_k)]^{-1} \nabla f(\mathbf{x}_k) . \quad (3.122)$$

Tomando a norma de ambos os lados, obtém-se a relação:

$$\begin{aligned} \| [\nabla^2 f(\mathbf{x}_k)]^{-1} \nabla^2 f(\mathbf{x}_k) \mathbf{d}_k \| &\leq \| [\nabla^2 f(\mathbf{x}_k)]^{-1} \| \| \nabla^2 f(\mathbf{x}_k) \| \| \mathbf{d}_k \|, \text{ ou} \\ \| \mathbf{d}_k \| &\leq C \| \mathbf{d}_k \| \quad (C \text{ é o número de condição da Hessiana}). \end{aligned} \quad (3.123)$$

Reescrevendo a condição de ângulo e considerando (3.123), tem-se:

$$\nabla f(\mathbf{x}_k)^T \mathbf{d}_k \leq -\theta \| \nabla f(\mathbf{x}_k) \| \| \mathbf{d}_k \| \leq -\theta C \| \nabla f(\mathbf{x}_k) \| \| \mathbf{d}_k \|^2$$

$$\theta \leq - \frac{1}{C} \frac{\nabla f(\mathbf{x}_k)^T \mathbf{d}_k}{\| \nabla f(\mathbf{x}_k) \| \| \mathbf{d}_k \|^2} . \quad (3.124)$$

Como $\theta \in (0,1)$, utiliza-se apenas o módulo do lado direito, resultando:

$$\theta \leq \left| -\frac{1}{C} \frac{\nabla f(\mathbf{x}_k)^T \mathbf{d}_k}{\|\nabla f(\mathbf{x}_k)\| \|\mathbf{d}_k\|} \right| \leq \frac{1}{C}. \quad (3.125)$$

O limite superior para θ é o inverso do número de condição da matriz Hessiana.

A decomposição de valor singular da matriz Hessiana forneceria de imediato o número de condição e o maior autovalor de $\nabla^2 f(\mathbf{x}_k)$. Mas o custo computacional em se realizar a decomposição de Cholesky no passo (4) do algoritmo 3.9 e depois realizar uma nova decomposição de valor singular de $\nabla^2 f(\mathbf{x}_k)$ tornaria o algoritmo inviável.

Tanto o maior autovalor de $\nabla^2 f(\mathbf{x}_k)$, quanto o número de condição, podem ser estimados aproximadamente utilizando os círculos de Gerschgorin.

Os limites superiores e inferiores dos autovalores de uma matriz simétrica $\mathbf{A} \in \mathbb{R}^{n \times n}$, utilizando os círculos de Gerschgorin são (DENNIS JR e SCHNABEL, 1983, p. 60) :

$$\min_{1 \leq i \leq n} \lambda_i \geq \min_{1 \leq i \leq n} \left\{ a_{ii} - \sum_{\substack{j=1 \\ j \neq i}}^n |a_{ij}| \right\}; \quad (3.126)$$

$$\max_{1 \leq k \leq n} \lambda_k \leq \max_{1 \leq k \leq n} \left\{ a_{kk} + \sum_{\substack{j=1 \\ j \neq k}}^n |a_{kj}| \right\}. \quad (3.127)$$

A estimativa do maior autovalor de $\nabla^2 f(\mathbf{x}_k)$ pode se tirado diretamente de (3.127), e o número de condição pode ser estimado por:

$$C \approx \frac{\lambda_{\text{MaxGer}}}{|\lambda_{\text{MinGer}}|}, \text{ para } \lambda_{\text{MinGer}} \neq 0. \quad (3.128)$$

A condição de Armijo, $f(\mathbf{x}_k + \lambda \mathbf{d}_k) \leq f(\mathbf{x}_k) + \alpha \lambda \nabla f(\mathbf{x}_k)^T \mathbf{d}_k$, deve ser satisfeita na fase (7) do algoritmo 3.9 para que haja decréscimo suficiente na função objetivo. A primeira tentativa é com $\lambda = 1$ e se a condição não for cumprida, deve-se escolher $\lambda \in$

$[0,1\lambda, 0,9\lambda]$, implicando numa busca no sentido retrógrado ao longo da linha definida pela direção de \mathbf{d}_k .

Uma implementação mais eficiente do que adotar o valor de $\lambda \in [0,1\lambda, 0,9\lambda]$ é utilizar modelos quadráticos e cúbicos. O modelo quadrático é utilizado na primeira busca do valor do novo λ e a partir de uma segunda busca, utiliza-se o modelo cúbico por melhor se adequar quando a função $f(\mathbf{x})$ tem curvatura negativa.

A busca do valor de λ de modo a satisfazer a condição de Armijo é denominada em inglês por *backtracking*. A descrição detalhada da busca com *backtracking* usando os modelos quadráticos e cúbicos, assim como o algoritmo, são encontrados em DENNIS JR e SCHNABEL (1983, p. 126, 325).

A implementação realizada nesse trabalho utilizando busca por *backtracking* aplica o algoritmo de Dennis e Schnabel.

Os passos 2, 3 e 4 do algoritmo 3.9 podem ser substituídos por um único, se houvesse a garantia de que a matriz \mathbf{B}_k , de alguma forma fosse definida positiva. De fato, esta possibilidade é obtida com a decomposição de Choleky modificada, descrita resumidamente a seguir.

Se $\mathbf{A} \in \mathbf{R}^{n \times n}$ é uma matriz simétrica e não necessariamente definida positiva, a decomposição de Cholesky modificada (\mathbf{LL}^T ou \mathbf{LDL}^T) é feita sobre uma matriz simétrica e definida positiva $\mathbf{A}+\mathbf{E}$ de maneira a satisfazer as seguintes metas (SCHNABEL; ESKOW, 1999, p.1135):

- i) se \mathbf{A} é sem perigo definida positiva, $\mathbf{E}=\mathbf{0}$;
- ii) se \mathbf{A} é indefinida $\|\mathbf{E}\|_\infty$ não é muito maior do que o maior autovalor negativo de \mathbf{A} ;
- iii) $\mathbf{A}+\mathbf{E}$ é razoavelmente bem condicionada;
- iv) o custo em obtê-la é somente um pequeno múltiplo de n^2 a mais do que o custo $O(n^3)$ da decomposição padrão.

A matriz \mathbf{E} , geralmente é diagonal, mas segundo SCHNABEL e ESKOW

(1999, p. 1138), CHENG e HIGHAM (1998)¹⁷, propuseram adicionar em \mathbf{A} uma matriz \mathbf{E} não diagonal, obtendo a decomposição de Cholesky modificada na forma \mathbf{LBL}^T , com \mathbf{B} bloco diagonal. Mas, segundo SCHNABEL e ESKOW (1999, p. 1138), os próprios autores do método, comentam que os resultados não são bons quando número de condição de \mathbf{LL}^T é alto, ao passo que, com \mathbf{E} na forma diagonal, SCHNABEL e ESKOW (1999, p. 1138), afirmam que os resultados têm sido bastante satisfatórios.

Os algoritmos da decomposição de Cholesky modificada, mais empregados da área de otimização são devido a GILL, MURRAY e WRIGHT (1981, p.111), DENNIS JR e SCHNABEL (1983, p. 318) e SCHNABEL e ESKOW (1999, p. 1140).

Para o entendimento do princípio da decomposição de Cholesky modificada, considere que no estágio j , ainda resta uma submatriz \mathbf{A}_j ($(n-j+1) \times (n-j+1)$) ($\mathbf{A}_1 = \mathbf{A}$) para completar o processo, \mathbf{A}_j tendo a seguinte forma:

$$\mathbf{A}_j = \begin{bmatrix} \alpha_j & \mathbf{a}_j^T \\ \mathbf{a}_j & \hat{\mathbf{A}}_j \end{bmatrix}, \quad (3.129)$$

onde: α_j é o elemento pivô; \mathbf{a}_j é um vetor contendo os elementos atuais da coluna j abaixo da diagonal de \mathbf{A}_j ; e $\hat{\mathbf{A}}_j$ é uma submatriz com dimensão $(n-j \times n-j)$.

Os elementos da matriz \mathbf{L} são calculados por:

$$\mathbf{L}_{jj} = \sqrt{\alpha_j + \delta_j}, \quad (3.130)$$

$$\mathbf{L}_{ij} = \frac{(\mathbf{a}_j)_i}{\mathbf{L}_{jj}}, \quad i = j+1, \dots, n, \quad (3.131)$$

$$\mathbf{A}_{j+1} = \hat{\mathbf{A}}_j - \frac{\mathbf{a}_j \mathbf{a}_j^T}{\alpha_j + \delta_j}. \quad (3.132)$$

A idéia por trás da decomposição de Cholesky modificada é encontrar um

¹⁷ CHENG, S.H.; HIGHAM, N.J. A modified cholesky algorithm based on a symmetric indefinite factorization. **SIAM J. Matrix Anal. Appl.**, 19, 1998, p. 1097-1110.

valor positivo para δ_j (elemento da diagonal de E) a fim de que se cumpra as 4 metas estabelecidas anteriormente. Isto não é tão simples, porque a escolha de $\delta > \alpha_j$ não é satisfatória. Os detalhes sobre esta escolha são esclarecidos em GILL, MURRAY e WRIGHT (1981, p. 109) e SCHNABEL e ESKOW (1999, p. 1137).

O algoritmo implementado nesse trabalho foi o de SCHNABEL e ESKOW (1999, p. 1140), denominado de algoritmo revisado da decomposição de Cholesky modificada. O algoritmo é transcrito a seguir na forma em que foi publicado.

Algoritmo 3.10 (Algoritmo revisado da decomposição de Cholesky modificada).

Dado $A \in R^{n \times n}$ simétrica (armazenada na triangular inferior) e τ , $\bar{\tau}$, μ (exemplo, $\tau = (\text{macheps})^{\frac{1}{3}}$, $\bar{\tau} = (\text{macheps})^{\frac{2}{3}}$, $\mu = 0,1$), encontrar a decomposição LL^T de $A+E$, $E \geq 0$.

```

phaseone := true
 $\gamma := \max_{1 \leq i \leq n} \{ |A_{ii}| \}$ 
j := 1
(* Phase one, A potentially positive definite *)
While j ≤ n and phaseone = true do
  if  $\max_{j \leq i \leq n} \{ A_{ij} \} < \bar{\tau} \gamma$  or  $\min_{j \leq i \leq n} \{ A_{ij} \} < -\mu (\max_{j \leq i \leq n} \{ A_{ij} \})$ 
    then phaseone := true (* go to phase two *)
  else
    (* Pivot on maximum diagonal of remaining submatrix *)
    i := index of  $\max_{j \leq i \leq n} \{ A_{ii} \}$ 
    if  $i \neq j$ , switch rows and columns of i and j of A
    if  $\min_{j+1 \leq i \leq n} \{ A_{ii} - A_{ij}^2 / A_{jj} \} < -\mu \gamma$ 
      then phaseone := false (* go to phase two *)
      else (* perform jth iteration of factorization *)
         $L_{jj} = \sqrt{A_{jj}}$  (*  $L_{jj}$  overwrites  $A_{jj}$  *)
        For i := j+1 to n do
           $L_{ij} := A_{ij} / L_{jj}$  (*  $L_{ij}$  overwrites  $A_{ij}$  *)
          For k := j+1 to i do
             $A_{ik} = A_{ik} - L_{ij} L_{kj}$ 
          j := j+1
    (* end phase one *)

(* Phase two, A not positive definite *)
if phaseone = false and j = n, then
   $\delta (* = E_{nn} *) := -A_{nn} + \max \{ \tau(-A_{nn}) / (1-\tau), \bar{\tau} \gamma \}$ 
   $A_{nn} := A_{nn} + \delta$ 
   $L_{nn} = \sqrt{A_{nn}}$ 
if phaseone = false and j < n, then
  k := j - 1 (* k = number of iterations performed in phase one *)
  (* Calculate lower Gerschgorin bounds of  $A_{k+1}$  *)
  For i := k+1 to n do

```

```

       $g_i := A_{ii} - \sum_{j=k+1}^{i-1} |A_{ij}| - \sum_{j=k+1}^n |A_{ji}|$ 
      (* Modified Cholesky Decomposition*)
For j:=k+1 to n-2 do
      (* Pivot on maximum lower Gerschgorin bound estimate*)
      i:= index of  $\max_{j \leq i \leq n} \{g_i\}$ 
      if  $i \neq j$ , switch rows and columns of i and j of A
      (* Calculate  $E_{ij}$  and add to diagonal*)
       $norm_j := \sum_{i=j+1}^n |A_{ij}|$ 
       $\delta (=E_{nn}) := \max\{0, -A_{jj} + \max\{norm_j, \bar{\tau} \gamma\}, \delta_{prev}\}$ 
      if  $\delta > 0$ , then
           $A_{jj} := A_{jj} + \delta$ 
           $\delta_{prev} = \delta$  (* $\delta_{prev}$  will contain  $\|E\|_{\infty}$ *)
      (* Update Gerschgorin bound estimates*)
      if  $A_{ij} \neq norm_j$ , then
           $temp := 1 - norm_j / A_{ij}$ 
          for i := j + 1 to n do
               $g_i := g_i + |A_{ij}| * temp$ 
      (* Perform jth iteration of factorization *)
      same code as in phase one

(*Final 2x2 submatrix*)
 $\lambda_{lo}, \lambda_{hi} := \text{eigenvalues of } \begin{bmatrix} A_{n-1,n-1} & A_{n,n-1} \\ A_{n,n-1} & A_{n,n} \end{bmatrix}$ 
 $\delta := \max\{0, -\lambda_{lo} + \max\{\tau(\lambda_{hi} - \lambda_{lo})/(1-\tau), \bar{\tau} \gamma\}, \delta_{prev}\}$ 
if  $\delta > 0$ , then
     $A_{n-1,n-1} := A_{n-1,n-1} + \delta$ 
     $A_{n,n} := A_{n,n} + \delta$ 
     $\delta_{prev} = \delta$ 
     $L_{n-1,n-1} := \sqrt{A_{n-1,n-1}}$  (*overwrites  $A_{n-1,n-1}$ *)
     $L_{n,n-1} := A_{n,n-1} / L_{n-1,n-1}$  (*overwrites  $A_{n,n-1}$ *)
     $L_{n,n} := (A_{n,n} - L_{n,n-1}^2)^{1/2}$  (*overwrites  $A_{n,n}$ *)
(* End phase two*)

```

No algoritmo são realizadas algumas permutações de linhas e colunas. Com estas permutações o resultado da decomposição pode ser representado por:

$$A + P^T E P = P^T L L^T P, \text{ ou} \quad (3.133)$$

$$P A P^T + E = L L^T, \quad (3.134)$$

onde:

P : é uma matriz ortogonal de permutações.

Obs.: Agradecemos aos autores Dr. Robert B. Schanabel e Dra. Elizabeth Eskow, pela especial atenção de nos ter enviado a codificação em Fortran da decomposição.

Utilizando a decomposição de Cholesky modificada, o algoritmo 3.9 pode ser reformulado com a seguinte estratégia.

Algoritmo 3.11 (Newton modificado globalizado com decomposição de Cholesky modificada). Dados $\alpha \in (0,1)$, $\beta > 0$, $\theta \in (0,1)$ e $\mathbf{x}_k \in \mathbb{R}^n$

1. Se $\nabla f(\mathbf{x}_k) = \mathbf{0}$, parar.
2. Realizar a decomposição de Cholesky modificada de $\nabla^2 f(\mathbf{x}_k)$, obtendo:

$$\mathbf{B}_k = \nabla^2 f(\mathbf{x}_k) + \mathbf{P}^T \mathbf{E} \mathbf{P} = \mathbf{P}^T \mathbf{L} \mathbf{L}^T \mathbf{P},$$

resolver

$$\mathbf{L} \mathbf{z} = -\mathbf{P} \nabla f(\mathbf{x}_k)$$

$$\mathbf{L}^T \mathbf{y} = \mathbf{z}$$

$$\mathbf{d}_k = \mathbf{P}^T \mathbf{y} \quad (\mathbf{d}_k \text{ é com certeza uma direção de descida})$$

$$\text{Faça } \mu = \|\mathbf{E}\|_\infty = \max \{\text{diag}(\mathbf{E})\}$$

3. Enquanto $\nabla f(\mathbf{x}_k)^T \mathbf{d}_k > -\theta \|\nabla f(\mathbf{x}_k)\| \|\mathbf{d}_k\|$, fazer $\mu \leftarrow \max \{2\mu, 10\}$ e resolva $[\mathbf{B}_k + \mu \mathbf{I}] \mathbf{d}_k = -\nabla f(\mathbf{x}_k)$. Obs.: Como $[\mathbf{B}_k + \mu \mathbf{I}]$ já é definida positiva, pode-se usar a decomposição de Cholesky padrão ou apenas a primeira fase do algoritmo 3.10 da decomposição de Cholesky modificada.

4. Se $\|\mathbf{d}_k\| < \beta \|\nabla f(\mathbf{x}_k)\|$, corrigir:

$$\mathbf{d}_k \leftarrow \beta \frac{\mathbf{d}_k}{\|\mathbf{d}_k\|} \|\nabla f(\mathbf{x}_k)\|$$

5. Obter λ por *backtracking* usando modelos quadráticos e cúbicos, de modo a satisfazer a condição de Armijo

$$f(\mathbf{x}_k + \lambda \mathbf{d}_k) \leq f(\mathbf{x}_k) + \alpha \lambda \nabla f(\mathbf{x}_k)^T \mathbf{d}_k,$$

definir

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \lambda \mathbf{d}_k$$

e voltar para (1).

3.1.7.5 Método de Gauss-Newton modificado globalizado

O algoritmo 3.11 anteriormente apresentado, com a substituição da matriz Hessiana ($\nabla^2 f(\mathbf{x}_k)$) pela matriz com as informações de primeira ordem do problema de mínimos quadrados não-linear (equações normais $[\mathbf{J}(\mathbf{x}_k)^T \mathbf{Q}_{\text{lb}}^{-1} \mathbf{J}(\mathbf{x}_k)]$), será chamado Gauss-Newton modificado globalizado. Note que além de incorporar as características do método de Levenberg-Marquardt, equação (3.116), o algoritmo também incorpora as condições de globalização (fig. 3.5).

O algoritmo adaptado com a notação do problema de mínimos quadrados não-linear é descrito a seguir.

Algoritmo 3.12 (Gauss-Newton modificado globalizado). Dados $\alpha \in (0,1)$, $\beta > 0$, $\theta \in (0,1)$ e $\mathbf{x}_k \in \mathbb{R}^n$

1. Se $\|\nabla f(\mathbf{x}_k)\| = \|\mathbf{J}(\mathbf{x}_k)^T \mathbf{Q}_{\text{lb}}^{-1} \mathbf{r}(\mathbf{x}_k)\| < \text{tol}_{\text{grad}}$, parar.
2. Realizar a decomposição de Cholesky modificada de $[\mathbf{J}(\mathbf{x}_k)^T \mathbf{Q}_{\text{lb}}^{-1} \mathbf{J}(\mathbf{x}_k)]$ obtendo:

$$\mathbf{B}_k = [\mathbf{J}(\mathbf{x}_k)^T \mathbf{Q}_{\text{lb}}^{-1} \mathbf{J}(\mathbf{x}_k)] + \mathbf{P}^T \mathbf{E} \mathbf{P} = \mathbf{P}^T \mathbf{L} \mathbf{L}^T \mathbf{P},$$
 resolver

$$\mathbf{L} \mathbf{z} = -\mathbf{P} \mathbf{J}(\mathbf{x}_k)^T \mathbf{Q}_{\text{lb}}^{-1} \mathbf{r}(\mathbf{x}_k); \quad \mathbf{L}^T \mathbf{y} = \mathbf{z}$$

$$\mathbf{d}_k = \mathbf{P}^T \mathbf{y} \quad (\mathbf{d}_k \text{ é com certeza uma direção de descida})$$
 Faça $\mu = \|\mathbf{E}\|_\infty = \max \{\text{diag}(\mathbf{E})\}$
3. Enquanto $[\mathbf{J}(\mathbf{x}_k)^T \mathbf{Q}_{\text{lb}}^{-1} \mathbf{r}(\mathbf{x}_k)]^T \mathbf{d}_k > -\theta \|\mathbf{J}(\mathbf{x}_k)^T \mathbf{Q}_{\text{lb}}^{-1} \mathbf{r}(\mathbf{x}_k)\| \|\mathbf{d}_k\|$, faça $\mu \leftarrow \max \{2\mu, 10\}$ e resolva

$$[\mathbf{B}_k + \mu \mathbf{I}] \mathbf{d}_k = -\mathbf{J}(\mathbf{x}_k)^T \mathbf{Q}_{\text{lb}}^{-1} \mathbf{r}(\mathbf{x}_k)$$
 Obs.: Como $[\mathbf{B}_k + \mu \mathbf{I}]$ já é definida positiva, pode-se usar a decomposição de Cholesky padrão ou apenas a primeira fase do algoritmo 3.10 da decomposição de Cholesky modificada.
4. Se $\|\mathbf{d}_k\| < \beta \|\mathbf{J}(\mathbf{x}_k)^T \mathbf{Q}_{\text{lb}}^{-1} \mathbf{r}(\mathbf{x}_k)\|$, corrigir:

$$\mathbf{d}_k \leftarrow \beta \frac{\mathbf{d}_k}{\|\mathbf{d}_k\|} \|\mathbf{J}(\mathbf{x}_k)^T \mathbf{Q}_{\text{lb}}^{-1} \mathbf{r}(\mathbf{x}_k)\|$$
5. Obter λ por *backtracking* usando modelos quadráticos e cúbicos, de modo a satisfazer a condição de Armijo

$$f(\mathbf{x}_k + \lambda \mathbf{d}_k) \leq f(\mathbf{x}_k) + \alpha \lambda [\mathbf{J}(\mathbf{x}_k)^T \mathbf{Q}_{Lb}^{-1} \mathbf{r}(\mathbf{x}_k)]^T \mathbf{d}_k ,$$

definir

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \lambda \mathbf{d}_k$$

e voltar para (1).

3.2 FUNDAMENTOS ESTATÍSTICOS

Qualquer pessoa que se aventure a realizar várias medidas de uma grandeza, vai perceber, por mais que insista, que as medidas não vão se repetir todas as vezes. Esta propriedade aleatória das observações, certamente foi percebida até mesmo pelos antigos egípcios nos levantamentos topográficos do vale do Nilo, e que depois serviram a Eratóstenes (nascido em 276 a.c.) para determinar pela primeira vez, o diâmetro da Terra com um erro inferior a 1000 km¹⁸.

Durante os dois últimos séculos, as propriedades aleatórias das observações passaram a ser estudadas sob um ponto de visto estatístico, abordados pela teoria das probabilidades. As observações começaram então a ser consideradas realizações de variáveis aleatórias.

A seguir serão descritas algumas definições (na linguagem estatística) da teoria das probabilidades e que são empregadas no ajustamento.

Espaço amostral, Ω , de um experimento realizado sob condições fixas, é o conjunto de todos os resultados possíveis do experimento, entendendo-se por resultado possível todo resultado elementar e indivisível do experimento.

Evento: é qualquer subconjunto do espaço amostral Ω , incluindo o evento impossível Φ , e o evento certo Ω .

¹⁸ Um resumo histórico com os grandes nomes da Geodésia é encontrado em GEMAEL (1977, p. 6.2).

σ -Álgebra, \mathcal{A} , de subconjuntos do conjunto não-vazio Ω é a classe de subconjuntos de Ω satisfazendo as propriedades:

- i) $\Omega \in \mathcal{A}$;
- ii) se $A \in \mathcal{A}$, então $A^c \in \mathcal{A}$ (A^c denota o complemento de A);
- iii) se $A_i, i=1,2,\dots \in \mathcal{A}$ então $\bigcup_{i=1}^{\infty} A_i \in \mathcal{A}$.

Definição Axiomática de Probabilidade (Kolmogorov): probabilidade ou medida de probabilidade na σ -Álgebra \mathcal{A} , é uma função P definida em \mathcal{A} e que satisfaz aos axiomas:

- i) $0 \leq P(A) \leq 1$, para todo evento A associado a Ω ;
- ii) $P(\Omega)=1$;
- iii) se A e $B \in \mathcal{A}$ e são mutuamente exclusivos, então $P(A \cup B) = P(A) + P(B)$;
- iv) se $A_i, i=1,2,\dots \in \mathcal{A}$ e são mutuamente exclusivos, então $P(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i)$.

Variável Aleatória: é uma função que associa números reais aos eventos de um espaço amostral e para cada qual é associada uma probabilidade. Uma variável aleatória (X) pode ser contínua ou discreta. Será discreta quando o número de valores possíveis de X for finito ou infinito enumerável, e será contínua se o conjunto de valores que X pode assumir for um intervalo ou uma reunião de intervalos.

3.2.1 Esperança Matemática

A esperança matemática de uma variável aleatória discreta (X), se ela existe, é calculada pelo somatório dos valores x_i de X multiplicado pelas correspondentes probabilidades $P(x_i)$. Geralmente é denotada por $E\{X\}$ e corresponde ao valor médio μ_x da variável

$$E\{X\} = \mu_x = \sum x_i P(x_i), \quad i = 1, 2, 3, \dots \quad (3.135)$$

Se a variável aleatória (X) é contínua no intervalo $(-\infty, +\infty)$, então:

$$E\{X\} = \mu_X = \int_{-\infty}^{+\infty} x f(x) dx, \quad (3.136)$$

onde: $f(x)$ é a função densidade de probabilidade de X . A relação (3.136) pode ser estendida a uma forma mais geral quando se considera a esperança de uma função $g(X)$ cuja função densidade é $f(x)$, nesse caso:

$$E\{g(X)\} = \int_{-\infty}^{+\infty} g(x) f(x) dx. \quad (3.137)$$

A expressão (3.137) pode ser estendida a outras dimensões. Por exemplo, se a variável aleatória bidimensional $\mathbf{X} = (X_1, X_2)$ com correspondente função densidade conjunta $f(x_1, x_2)$ contínua, então para qualquer função $g(X_1, X_2)$ a esperança é definida como:

$$E\{g(X_1, X_2)\} = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} g(x_1, x_2) f(x_1, x_2) dx_1 dx_2. \quad (3.138)$$

De maneira semelhante a definição se estende ao caso n -dimensional

$$E\{g(X_1, X_2, \dots, X_n)\} = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} g(x_1, x_2, \dots, x_n) f(x_1, x_2, \dots, x_n) dx_1 dx_2, \dots, dx_n. \quad (3.139)$$

Com base nas definições (3.135) a (3.139), obtém-se as propriedades da esperança matemática:

- i) $E\{c\} = c$, onde c é uma constante ;
- ii) $E\{c X\} = c E\{X\}$;
- iii) $E\{X+Y\} = E\{X\}+E\{Y\}$;
- iv) $E\{XY\} = E\{X\}E\{Y\}$, se e somente se, X e Y são variáveis aleatórias independentes .

3.2.2 Momentos, Variâncias e Covariâncias

Momento estatístico de ordem k de uma variável aleatória X é definido como

a esperança desta variável elevada a k-ésima potência. Se a variável for discreta, escreve-se:

$$m_k = E\{X^k\} = \sum_{i=1}^n x_i^k P(x_i), \quad (3.140)$$

e, se a variável aleatória for contínua, tem-se:

$$m_k = E\{X^k\} = \int_{-\infty}^{+\infty} x^k f(x) dx. \quad (3.141)$$

O momento de ordem k centrado em relação a uma constante (c), é definido como:

$$m_k = E\{(X - c)^k\}. \quad (3.142)$$

Quando c na equação (3.142) for $c = E\{X\} = \mu_x$, diz-se que o momento é central ou centrado com relação a média

$$m_k = E\{(X - E\{X\})^k\}. \quad (3.143)$$

Como caso particular, fazendo $k=2$ na equação (3.143), tem-se o momento central de segunda ordem, que é a variância da variável aleatória X

$$m_2 = \sigma_x^2 = E\{(X - \mu_x)^2\}. \quad (3.144)$$

Fazendo nas equações (3.140) ou (3.141), $k=1$, obtém-se o valor médio da variável aleatória, conforme seja, discreta ou contínua.

Considere agora uma variável aleatória bidimensional $X=(X_1, X_2)$. O momento central de ordem $(l+n)$ dessa variável aleatória é definido como:

$$m_{ln} = E\{(X_1 - \mu_{x1})^l (X_2 - \mu_{x2})^n\}. \quad (3.145)$$

Para alguns momentos centrais particulares, tais como m_{20} , m_{02} , m_{11} , tem-se:

$$m_{20} = E\{(X_1 - \mu_{x1})^2\} = \sigma_{x1}^2 ; \quad (\text{variância de } X_1)$$

$$m_{02} = E\{(X_2 - \mu_{x2})^2\} = \sigma_{x2}^2 ; \quad (\text{variância de } X_2)$$

$$m_{11} = E\{(X_1 - \mu_{x1})(X_2 - \mu_{x2})\} = \sigma_{x1x2} \quad (\text{covariância entre } X_1 \text{ e } X_2).$$

No caso geral de uma variável aleatória n-dimensional \mathbf{X} , todos os momentos centrais de segunda ordem podem ser combinados e arranjados dentro de uma matriz \mathbf{M}_x ,

$$\mathbf{M}_x = \begin{bmatrix} m_{x_1x_1} & m_{x_1x_2} & \cdot & \cdot & \cdot & m_{x_1x_n} \\ m_{x_2x_1} & m_{x_2x_2} & \cdot & \cdot & \cdot & m_{x_2x_n} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ m_{x_nx_1} & m_{x_nx_2} & \cdot & \cdot & \cdot & m_{x_nx_n} \end{bmatrix} \quad (3.146)$$

A matriz \mathbf{M}_x de momentos centrais de segunda ordem é exatamente a matriz de variâncias e covariâncias da variável aleatória n-dimensional \mathbf{X} , cuja representação é:

$$\Sigma_x = \begin{bmatrix} \sigma_{x_1}^2 & \sigma_{x_1x_2} & \cdot & \cdot & \cdot & \sigma_{x_1x_n} \\ \sigma_{x_2x_1} & \sigma_{x_2}^2 & \cdot & \cdot & \cdot & \sigma_{x_2x_n} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \sigma_{x_nx_1} & \sigma_{x_nx_2} & \cdot & \cdot & \cdot & \sigma_{x_n}^2 \end{bmatrix} \quad (3.147)$$

A matriz Σ_x é conhecida por matriz variância-covariância ou simplesmente matriz de covariância da variável aleatória n-dimensional \mathbf{X} .

Para maiores detalhes a respeito das seções 3.2.1 e 3.2.2 consultar, por exemplo: MIKAIL e ACKERMAN (1976, p. 16-21).

3.2.3 Propagação das Covariâncias

Considere o modelo afim

$$\mathbf{Y} = \mathbf{G}\mathbf{X} + \mathbf{C} \quad (3.148)$$

onde:

\mathbf{X} : é uma variável aleatória n-dimensional;

$\mathbf{G} \in R^{m \times n}$: matriz de coeficientes (constantes);

$\mathbf{C} \in R^m$: vetor de constantes;

\mathbf{Y} : é uma variável aleatória m-dimensional.

Obs.: Com o modelo (3.148), se \mathbf{X} é variável aleatória, conseqüentemente \mathbf{Y} também será.

Expressando o momento central de segunda ordem de \mathbf{Y}

$$\Sigma_Y = E\{(\mathbf{Y} - \mu_Y)(\mathbf{Y} - \mu_Y)^T\}. \quad (3.149)$$

Lembrando das propriedades da esperança matemática

$$\mu_Y = E\{\mathbf{Y}\} = E\{\mathbf{G}\mathbf{X} + \mathbf{C}\} = \mathbf{G}\mu_X + \mathbf{C}, \quad (3.150)$$

e substituindo (3.150) e (3.148) em (3.149), resulta:

$$\Sigma_Y = \mathbf{G}E\{(\mathbf{X} - \mu_X)(\mathbf{X} - \mu_X)^T\}\mathbf{G}^T. \quad (3.151)$$

Observando que $\Sigma_X = E\{(\mathbf{X} - \mu_X)(\mathbf{X} - \mu_X)^T\}$ é a matriz variância-covariância da variável aleatória (\mathbf{X}), a equação (3.151), torna-se:

$$\Sigma_Y = \mathbf{G} \Sigma_X \mathbf{G}^T. \quad (3.152)$$

A equação (3.152), expressa a conhecida lei de propagação das covariâncias para funções lineares (modelo afim 3.148).

Quando $Y=F(X)$, sendo $F:R^n \rightarrow R^m$ não-linear, a matriz G em (3.152) é substituída pela matriz Jacobiano $J(X)=\frac{\partial F}{\partial X}/_{x_0}$, sendo X_0 um vetor de valores aproximados. Em Geodésia costuma-se usar a notação D em vez de $J(X)$. Assim, a lei de propagação das covariâncias para funções não-lineares é representada por:

$$\Sigma_Y = D \Sigma_X D^T. \quad (3.153)$$

3.2.4 Matriz dos Pesos para as Observáveis GPS

Ao longo de todo o desenvolvimento da seção 3.1 e respectivas subseções, introduziu-se uma matriz simétrica e definida positiva denotada por Q_L^{-1} , chamada matriz dos pesos. Nessa seção são esclarecidos alguns detalhes relacionados a ela, quando utilizada no ajustamento das observáveis GPS.

De acordo com MIKAIL e GRACIE (1981, p.66), o peso de uma simples medida¹⁹ pode ser definido como o inverso de sua variância:

$$p = \frac{k}{\sigma^2}, \quad (3.154)$$

onde: k é uma constante de proporcionalidade.

Para uma medida (observação) com peso unitário ($p = 1$) a relação (3.154), torna-se:

$$k = \sigma^2. \quad (3.155)$$

Assim, a constante de proporcionalidade k para a medida com peso unitário, passa a ser a variância da medida (na verdade é a variância da variável aleatória da

¹⁹ Uma medida é uma realização da variável aleatória. Algumas medidas caracterizam uma amostra da variável aleatória enquanto que todas as medidas possíveis caracterizam a população que descreve a variável aleatória.

qual a medida foi realizada). Essa variância para a medida com peso unitário é denotada por σ_0^2 , sendo conhecida como variância de referência, fator de variância e variância da unidade de peso (ver MIKAIL; GRACIE, 1981, p. 67).

O peso de acordo com as relações anterior é então definido como:

$$p = \frac{\sigma_0^2}{\sigma^2}. \quad (3.156)$$

O inverso do peso (variância relativa), é conhecido como cofator ou coeficiente de peso, e denota-se:

$$q = \frac{\sigma^2}{\sigma_0^2}. \quad (3.157)$$

No caso da variável aleatória (\mathbf{X}), n -dimensional, as relações (3.156) e (3.157) podem ser generalizadas pelas seguintes formas:

$$\mathbf{P}_x = \sigma_0^2 \Sigma_x^{-1}; \quad (3.158)$$

$$\mathbf{Q}_x = \frac{1}{\sigma_0^2} \Sigma_x, \quad (3.159)$$

onde: \mathbf{P}_x é a matriz dos pesos; e \mathbf{Q}_x é a matriz cofatora.

Das equações (3.158) e (3.159), tira-se:

$$\mathbf{P}_x = \mathbf{Q}_x^{-1} \quad \text{ou} \quad \mathbf{Q}_x = \mathbf{P}_x^{-1}. \quad (3.160)$$

Uma variável aleatória n -dimensional é também chamada de vetor aleatório e, cada componente deste vetor é uma variável aleatória. Se duas variáveis aleatórias, componentes do vetor aleatório, forem independentes, a covariância entre elas será nula conforme mostra o resultado a seguir.

Sejam x_1 e x_2 duas variáveis aleatórias (unidimensionais) independentes, a covariância entre elas é dada por:

$$\sigma_{x_1 x_2} = E\{(x_1 - \mu_{x_1})(x_2 - \mu_{x_2})\}. \quad (3.161)$$

Desenvolvendo tem-se:

$$\begin{aligned} \sigma_{x_1 x_2} &= E\{x_1 x_2 - x_1 \mu_{x_2} - \mu_{x_1} x_2 + \mu_{x_1} \mu_{x_2}\} \\ &= E\{x_1 x_2\} - E\{x_1\} \mu_{x_2} - \mu_{x_1} E\{x_2\} + \mu_{x_1} \mu_{x_2} \\ &= E\{x_1 x_2\} - \mu_{x_1} \mu_{x_2}. \end{aligned} \quad (3.162)$$

Considerando a independência entre elas, a propriedade (iv) da esperança matemática pode ser aplicada, resultando:

$$\sigma_{x_1 x_2} = E\{x_1\} E\{x_2\} - \mu_{x_1} \mu_{x_2} = 0. \quad (3.163)$$

Desta forma, se todas as variáveis aleatórias componentes do vetor aleatório (\mathbf{X}) n-dimensional, forem independentes entre si, a matriz variância-covariância (Σ_x), será diagonal, e seus elementos não nulos serão as variâncias de cada variável aleatória.

Conforme foi visto no capítulo 2 sobre as equações de observação GPS, o receptor registra a pseudodistância receptor/satélite (P_R) (ver seção 2.6.1) e a fase da portadora (φ_R^s) (ver seção 2.6.2). Para cada satélite rastreado serão registradas, por época, essas medidas. Deste modo, a pseudodistância (variável aleatória) envolvendo o receptor R e o satélite S pode ser representada por P_R^s ; a pseudodistância (variável aleatória) envolvendo o receptor R e o satélite P, poder ser representada por P_R^p , e assim por diante. O mesmo pode ser dito para a fase da portadora (variável aleatória).

As variáveis aleatórias representadas pelas observáveis das pseudodistância e pelas fases das portadoras podem ser consideradas independentes entre si e entre épocas. As variâncias dessas variáveis aleatórias estão associadas ao ruído do receptor (ver seção 2.5.7). O ruído de cada observação pode ser interpretado como sendo o

desvio padrão²⁰ da correspondente variável aleatória (ver tab. 1.1 e tab. 2.6).

As variâncias de referência (σ_0^2), com base na tabela 1.1, podem ter os seguintes valores: $\sigma_{\sigma(C/A)}^2 = 4 \text{ m}^2$ (variância de referência para o código C/A), $\sigma_{\sigma(P)}^2 = 4 \text{ cm}^2$ (variância de referência para o código P), $\sigma_{\sigma(L_1)}^2 = \sigma_{\sigma(L_2)}^2 = 1 \text{ cm}^2$ (variância de referência para as fases das portadoras L_1 e L_2).

De acordo com a discussão acima e da definição (3.156), a matriz dos pesos ($\mathbf{P} = \mathbf{Q}_{\text{lb}}^{-1}$) para as equações de observação GPS (equação 2.63 e 2.79), será uma matriz identidade.

Matriz dos pesos para a simples diferença

As equações de observação da simples diferença entre receptores foram definidas pela equação (2.89) para os códigos e equação (2.90) para as fases.

Considerando, por exemplo, 5 satélites rastreados e usando medidas de fase, as 5 observações da simples diferença de fase para uma época, podem ser representadas por:

$$\begin{bmatrix} \phi_{AB}^1 \\ \phi_{AB}^2 \\ \phi_{AB}^3 \\ \phi_{AB}^4 \\ \phi_{AB}^5 \end{bmatrix} = \begin{bmatrix} 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 \end{bmatrix} \begin{bmatrix} \phi_B^1 \\ \phi_A^1 \\ \phi_B^2 \\ \phi_A^2 \\ \phi_B^3 \\ \phi_A^3 \\ \phi_B^4 \\ \phi_A^4 \\ \phi_B^5 \\ \phi_A^5 \end{bmatrix} \quad (3.164)$$

²⁰ O desvio padrão (σ) é definido com sendo a raiz quadrada positiva da variância (σ^2), ou seja, $\sigma = +\sqrt{\sigma^2}$.

O sobrescrito 1,2,...,5 denotam satélites, e os subscritos A e B denotam receptores. O vetor do lado direito de (3.164) é um vetor aleatório. De acordo com a Lei de propagação das covariâncias (equação 3.152), a matriz variância-covariância da simples diferença de fase é:

$$\Sigma_{\text{SDF}} = \mathbf{G} \Sigma_{\text{F}} \mathbf{G}^T, \text{ ou} \quad (3.165)$$

$$\Sigma_{\text{SDF}} = 2\sigma_{\phi}^2 \mathbf{I}_{(5 \times 5)}, \quad (3.166)$$

onde :

σ_{ϕ}^2 : é a variância das fases da portadora (as populações são idênticas).

A matriz dos pesos da simples diferença, para cada época, de acordo com as expressões (3.158) e (3.159), é:

$$\mathbf{P}_{\text{SDF}} = \mathbf{Q}_{\text{SDF}}^{-1} = \frac{\sigma_0^2}{2\sigma_{\phi}^2} \mathbf{I} = \frac{1}{2} \mathbf{I}. \quad (3.167)$$

Obs.: A matriz dos pesos para a simples diferença de pseudodistância é a mesma (3.167).

Matriz dos pesos para a dupla diferença

Considerando novamente os 5 satélites rastreados e fixando um deles para formar as duplas diferenças, por exemplo, o satélite 1, as 4 variáveis aleatórias da dupla diferença podem ser representadas por:

$$\begin{bmatrix} \phi_{AB}^{12} \\ \phi_{AB}^{13} \\ \phi_{AB}^{14} \\ \phi_{AB}^{15} \end{bmatrix} = \begin{bmatrix} 1 & -1 & -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 1 & 1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 1 & 0 & 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & -1 & 1 & 0 & 0 & 0 & 0 & 1 & -1 \end{bmatrix} \begin{bmatrix} \phi_B^2 \\ \phi_A^2 \\ \phi_B^1 \\ \phi_A^1 \\ \phi_B^3 \\ \phi_A^3 \\ \phi_B^4 \\ \phi_A^4 \\ \phi_B^5 \\ \phi_A^5 \end{bmatrix} \quad (3.167)$$

Novamente, aplicando a Lei de propagação das covariâncias para as variáveis aleatórias das duplas diferenças, tem-se:

$$\Sigma_{DDF} = \sigma_{\phi}^2 \begin{bmatrix} 4 & 2 & 2 & 2 \\ 2 & 4 & 2 & 2 \\ 2 & 2 & 4 & 2 \\ 2 & 2 & 2 & 4 \end{bmatrix} \quad (3.168)$$

A matriz dos pesos para a dupla diferença quando é fixado um satélite, seja fase ou código, terá a forma:

$$\mathbf{P}_{DDF} = \mathbf{Q}_{DDF}^{-1} = \frac{\sigma_0^2}{\sigma_{\phi}^2} \begin{bmatrix} 4 & 2 & 2 & 2 \\ 2 & 4 & 2 & 2 \\ 2 & 2 & 4 & 2 \\ 2 & 2 & 2 & 4 \end{bmatrix}^{-1} \quad (3.169)$$

A matriz dos pesos (\mathbf{P}_{DDF}), obtida da inversa da matriz cofatora (\mathbf{Q}_{DDF}) na equação (3.169), pode ser obtida diretamente por uma fórmula explícita em função do número de duplas diferenças independentes. Assim, se n for o número de duplas diferenças, a matriz dos pesos será (HOFMANN-WELLENHOF; LICHTENEGGER; COLLINS, 1994, p. 189):

$$\mathbf{P}_{\text{DDF}} = \mathbf{Q}_{\text{DDF}}^{-1} = \frac{\sigma_0^2}{\sigma_\phi^2} \begin{bmatrix} 4 & 2 & \cdots & 2 \\ 2 & 4 & \cdots & 2 \\ \vdots & \vdots & \ddots & \vdots \\ 2 & 2 & \cdots & 4 \end{bmatrix}^{-1} = \frac{\sigma_0^2}{\sigma_\phi^2} \frac{1}{2(n+1)} \begin{bmatrix} n & -1 & \cdots & -1 \\ -1 & n & \cdots & -1 \\ \vdots & \vdots & \ddots & \vdots \\ -1 & -1 & \cdots & n \end{bmatrix}. \quad (3.170)$$

Na formação da dupla diferença, não necessariamente é preciso usar um satélite de referência como foi na equação (3.167), pode-se, por exemplo, usar a formação sequencial $(\phi_{AB}^{12}, \phi_{AB}^{23}, \phi_{AB}^{34}, \phi_{AB}^{45}, \dots)$. Com esse tipo de formação, se o sinal de um satélite for perdido, duas equações serão afetadas, ao passo que na formação usando um satélite de referência $(\phi_{AB}^{12}, \phi_{AB}^{13}, \phi_{AB}^{14}, \phi_{AB}^{15}, \dots)$, se o sinal do satélite de referência for perdido, todas as equações serão afetadas, mas por outro lado, senão for o sinal do satélite de referência a ser perdido, apenas uma equação será afetada (mais detalhes sobre formação sequencial das DDF e de sua respectiva matriz variância-covariância, podem ser encontrados em MONICO, 1995, p. 45-47).

Para tentar evitar a perda do sinal do satélite de referência, o critério geralmente adotado é escolher como satélite de referência, aquele que estiver com maior ângulo de elevação.

As épocas de observações da dupla diferença podem ser consideradas independentes. Assim, por exemplo, se t épocas forem observadas, a matriz dos pesos para as duplas diferenças será bloco diagonal com a seguinte forma:

$$\mathbf{Q}_{\text{DDF}}^{-1} = \begin{bmatrix} \mathbf{Q}_{1(\text{DDF})}^{-1} & 0 & \cdots & 0 \\ 0 & \mathbf{Q}_{2(\text{DDF})}^{-1} & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & \mathbf{Q}_{t(\text{DDF})}^{-1} \end{bmatrix}, \quad (3.171)$$

sendo que cada bloco tem a estrutura dado pela equação (3.170).

Por ser diagonalmente dominante e ter posto completo, a matriz (3.170) e conseqüentemente a (3.171) são definidas positivas, admitindo assim, a decomposição

de Cholesky e, permitindo escrever o problema de mínimos quadrados linear conforme equação (3.13).

O procedimento para obter a matriz dos pesos para qualquer outra observável é análogo ao realizado anteriormente.

3.2.5 Caracterização Estatística para o Problema de Mínimos quadrados

Foi visto que a solução do problema de mínimos quadrados pode ser encontrada explicitamente com a expressão:

$$\mathbf{x}_a = (\mathbf{A}^T \mathbf{Q}_L^{-1} \mathbf{A})^{-1} \mathbf{A}^T \mathbf{Q}_L^{-1} \mathbf{L} \quad (3.172)$$

onde:

$\mathbf{L} \in \mathbf{R}^m$: é o vetor de observações;

$\mathbf{x}_a \in \mathbf{R}^n$: é o vetor de parâmetros do modelo $\mathbf{L}_a = \mathbf{F}(\mathbf{x}_a)$;

$\mathbf{Q}_L^{-1} \in \mathbf{R}^{m \times m}$: é a matriz dos pesos conforme esclarecido na seção anterior.

Se $\mathbf{F} : \mathbf{R}^n \rightarrow \mathbf{R}^m$ é não linear, foram descritos vários métodos na seção (3.1.7) para encontrar \mathbf{x}_a . Tomando uma iteração antes do método ter atingido a convergência, espera-se que a equação (3.172) possa representar adequadamente a solução (\mathbf{x}_a).

Aplicando a lei de propagação das covariâncias (equação 3.153) em (3.172), tem-se:

$$\begin{aligned} \Sigma_{\mathbf{x}_a} &= (\mathbf{A}^T \mathbf{Q}_L^{-1} \mathbf{A})^{-1} \mathbf{A}^T \mathbf{Q}_L^{-1} \Sigma_L \mathbf{Q}_L^{-1} \mathbf{A} (\mathbf{A}^T \mathbf{Q}_L^{-1} \mathbf{A})^{-1} \\ &= \sigma_o^2 (\mathbf{A}^T \mathbf{Q}_L^{-1} \mathbf{A})^{-1} \end{aligned} \quad (3.173)$$

A matriz variância-covariância dos parâmetros ajustados é obtida invertendo a matriz das equações normais ($\mathbf{A}^T \mathbf{Q}_L^{-1} \mathbf{A}$), justificando desta forma o algoritmo 3.2, e as equação (3.40), (3.47) e (3.60), comentadas anteriormente nos métodos de solução do

problema de mínimos quadrados linear.

No contexto estatístico, $\mathbf{x}_a = (\mathbf{A}^T \mathbf{Q}_L^{-1} \mathbf{A})^{-1} \mathbf{A}^T \mathbf{Q}_L^{-1} \mathbf{L}$, é uma variável aleatória n-dimensional. O vetor das observações é uma variável aleatória m-dimensional e as observações efetivadas correspondem a uma realização amostral da variável aleatória.

Quando a partir de uma amostra deseja-se conhecer algumas inferências sobre a população, o que se pode fazer é estimar valores para alguns parâmetros que identifiquem as características desta. A variável aleatória usada para a estimação é chamada estimador, enquanto que os valores numéricos obtidos por este estimador são chamados estimativas.

Desta forma, a equação (3.173) corresponde a um estimador das variâncias e covariâncias do vetor aleatório dos parâmetros ajustados. A equação (3.172) corresponde ao estimador de mínimos quadrados dos parâmetros do modelo $\mathbf{L}=\mathbf{F}(\mathbf{x})$.

Se o vetor dos resíduos aleatórios (\mathbf{V}), é proveniente de uma população de resíduos cuja média é zero, isto é, $E\{\mathbf{V}\}=\mathbf{0}$, o estimador de mínimos quadrados é um estimador não-tendencioso (não-viciado) dos parâmetros \mathbf{x} .

Reescrevendo a expressão do vetor dos resíduos

$$\mathbf{V}=\mathbf{Ax}-\mathbf{L}, \quad (3.174)$$

e aplicando a esperança matemática com a consideração acima, tem-se:

$$E\{\mathbf{V}\}=\mathbf{0}=\mathbf{AE}\{\mathbf{x}\}-E\{\mathbf{L}\} \Rightarrow E\{\mathbf{L}\}=\mathbf{AE}\{\mathbf{x}\}. \quad (3.175)$$

Aplicando a esperança em (3.172), tem-se:

$$E\{\mathbf{x}_a\} = (\mathbf{A}^T \mathbf{Q}_L^{-1} \mathbf{A})^{-1} \mathbf{A}^T \mathbf{Q}_L^{-1} E\{\mathbf{L}\}. \quad (3.176)$$

Substituindo (3.175) em (3.176), resulta:

$$E\{\mathbf{x}_a\} = E\{\mathbf{x}\} , \quad (3.177)$$

como $E\{\mathbf{x}\} = \mu_x$, então $E\{\mathbf{x}_a\} = \mu_x$, comprovando a não tendenciosidade das estimativas de \mathbf{x}_a para os parâmetros populacionais.

Uma outra qualidade desejável para um estimador não-tendencioso é que as variâncias de suas estimativas sejam mínimas. Com as hipóteses $E\{\mathbf{V}\} = \mathbf{0}$ e $\Sigma_L = \Sigma_V$ definida positiva, demonstra-se que o estimador de mínimos quadrados não-tendencioso, é também um estimador de variâncias mínimas (Prova: GARNÉS, 1998^(b), p.43), (ver também HAMILTON, 1964, p.143).

Se as observações são retiradas de uma população normalmente distribuída, o vetor dos resíduos é também normalmente distribuído e se $E\{\mathbf{V}\} = \mathbf{0}$, as estimativas dadas pelo estimador de mínimos quadrados são as de máxima probabilidade (MIKAIL; ACKERMAN, 1976, p.48).

Um último resultado fundamental relacionado às estimativas de mínimos quadrados é com respeito a variância de referência σ_0^2 . De acordo com a equação (3.155) e depois com as discussões sobre a formação da matriz dos pesos, a variância de referência corresponde (nesse caso) à variância populacional de onde as observações (amostras) foram extraídas.

Com a hipótese de que $E\{\mathbf{V}\} = \mathbf{0}$ pode-se encontrar um estimador não-tendencioso para σ_0^2 .

O vetor dos resíduos estimados pode ser representado por (ver equação 3.9):

$$\hat{\mathbf{V}} = \mathbf{A}\mathbf{x}_a - \mathbf{L} , \quad (3.178)$$

e a forma quadrática $(\hat{\mathbf{V}}^T \mathbf{Q}_L^{-1} \hat{\mathbf{V}})$, por:

$$\begin{aligned} \hat{\mathbf{V}}^T \mathbf{Q}_L^{-1} \hat{\mathbf{V}} &= \mathbf{L}^T \mathbf{Q}_L^{-1} \mathbf{L} - \mathbf{x}_a^T \mathbf{A}^T \mathbf{Q}_L^{-1} \mathbf{L} \\ &= \mathbf{L}^T \mathbf{Q}_L^{-1} \mathbf{L} - \mathbf{x}_a^T (\mathbf{A}^T \mathbf{Q}_L^{-1} \mathbf{A}) \mathbf{x}_a . \end{aligned} \quad (3.179)$$

Dividindo ambos os lados da equação (3.179) por σ_0^2 e lembrando de (3.158), tem-se:

$$\hat{\mathbf{V}}^T \Sigma_L^{-1} \hat{\mathbf{V}} = \mathbf{L}^T \Sigma_L^{-1} \mathbf{L} - \mathbf{x}_a^T (\mathbf{A}^T \Sigma_L^{-1} \mathbf{A}) \mathbf{x}_a. \quad (3.180)$$

Aplicando a esperança matemática em (3.180) e lembrando da propriedade do traço de matriz $\mathbf{y}^T \mathbf{B} \mathbf{y} = \text{tr}(\mathbf{y} \mathbf{y}^T \mathbf{B})$, com \mathbf{B} simétrica, vem:

$$E\{\hat{\mathbf{V}}^T \Sigma_L^{-1} \hat{\mathbf{V}}\} = E\{\text{tr}(\mathbf{L} \mathbf{L}^T \Sigma_L^{-1})\} - E\{\text{tr}(\mathbf{x}_a \mathbf{x}_a^T (\mathbf{A}^T \Sigma_L^{-1} \mathbf{A}))\}. \quad (3.181)$$

Usando a propriedade $E\{\text{tr}(\mathbf{B})\} = \text{tr}(E\{\mathbf{B}\})$, obtém-se:

$$E\{\hat{\mathbf{V}}^T \Sigma_L^{-1} \hat{\mathbf{V}}\} = \text{tr}(E\{\mathbf{L} \mathbf{L}^T\} \Sigma_L^{-1}) - \text{tr}(E\{\mathbf{x}_a \mathbf{x}_a^T\} (\mathbf{A}^T \Sigma_L^{-1} \mathbf{A})). \quad (3.182)$$

Da definição de matriz variância-convariância, tira-se:

$$E\{\mathbf{x}_a \mathbf{x}_a^T\} = \Sigma_{\mathbf{x}_a} + \mu_{\mathbf{x}_a} \mu_{\mathbf{x}_a}^T \quad (3.183)$$

$$E\{\mathbf{L} \mathbf{L}^T\} = \Sigma_L + \mu_L \mu_L^T \quad (3.184)$$

Substituindo ambas em (3.182), resulta:

$$E\{\hat{\mathbf{V}}^T \Sigma_L^{-1} \hat{\mathbf{V}}\} = \text{tr}(\mathbf{I}_m + \mu_L \mu_L^T \Sigma_L^{-1}) - \text{tr}(\mathbf{I}_n + \mu_{\mathbf{x}_a} \mu_{\mathbf{x}_a}^T (\mathbf{A}^T \Sigma_L^{-1} \mathbf{A})). \quad (3.185)$$

Usando a propriedade $\text{tr}(\mathbf{A} + \mathbf{B}) = \text{tr}(\mathbf{A}) + \text{tr}(\mathbf{B})$, tem-se que:

$$\begin{aligned} E\{\hat{\mathbf{V}}^T \Sigma_L^{-1} \hat{\mathbf{V}}\} &= m + \text{tr}(\mu_L \mu_L^T \Sigma_L^{-1}) - n - \text{tr}(\mu_{\mathbf{x}_a} \mu_{\mathbf{x}_a}^T (\mathbf{A}^T \Sigma_L^{-1} \mathbf{A})) \\ &= (m-n) + \mu_L^T \Sigma_L^{-1} \mu_L - \mu_{\mathbf{x}_a}^T (\mathbf{A}^T \Sigma_L^{-1} \mathbf{A}) \mu_{\mathbf{x}_a}. \end{aligned} \quad (3.186)$$

Com a hipótese $E\{\mathbf{V}\} = \mathbf{0}$, então $\mu_L = E\{\mathbf{L}\} = \mathbf{A} E\{\mathbf{L}\} = \mathbf{A} \mu_{\mathbf{x}_a}$. Substituindo μ_L em (3.186), resulta:

$$\begin{aligned} E\{\hat{\mathbf{V}}^T \Sigma_L^{-1} \hat{\mathbf{V}}\} &= (\mathbf{m}-\mathbf{n}) + \mu_x^T (\mathbf{A}^T \Sigma_L^{-1} \mathbf{A}) \mu_x - \mu_x^T (\mathbf{A}^T \Sigma_L^{-1} \mathbf{A}) \mu_x \\ &= (\mathbf{m}-\mathbf{n}). \end{aligned} \quad (3.187)$$

Substituindo $\Sigma_L^{-1} = \frac{1}{\sigma_0^2} \mathbf{Q}_L^{-1}$ em (3.187)

$$E\{\hat{\mathbf{V}}^T \mathbf{Q}_L^{-1} \hat{\mathbf{V}}\} = \sigma_0^2 (\mathbf{m}-\mathbf{n}). \quad (3.188)$$

Com esse resultado pode-se verificar facilmente que o estimador:

$$\hat{\sigma}_0^2 = \frac{\hat{\mathbf{V}}^T \mathbf{Q}_L^{-1} \hat{\mathbf{V}}}{(\mathbf{m}-\mathbf{n})}, \quad (3.189)$$

é um estimador não-tendencioso para σ_0^2 , isso porque:

$$E\{\hat{\sigma}_0^2\} = \frac{E\{\hat{\mathbf{V}}^T \mathbf{Q}_L^{-1} \hat{\mathbf{V}}\}}{(\mathbf{m}-\mathbf{n})} = \sigma_0^2. \quad (3.190)$$

O resultado demonstrado acima é um dos mais importantes para validação de um ajustamento por mínimos quadrados, pois, pode-se usar um teste de hipótese para verificar se a estimativa da variância de referência a posterior ($\hat{\sigma}_0^2$), a um certo nível de significância (α), é estatisticamente equivalente à variância de referência a priori (σ_0^2). Se a equivalência não for verificada é porque a hipótese $E\{\mathbf{V}\} = \mathbf{0}$ não está sendo verdadeira, isso contradiz alguns dos princípios fundamentais da distribuição de probabilidades das variáveis aleatórias envolvidas.

3.2.6 Testes de Hipóteses

Os testes estatísticos servem para verificar a hipótese de que uma determinada amostra (observações) foi, ou não, retirada de alguma população (especificada pelos parâmetros média e variância)

Geralmente duas hipóteses são testadas, a hipótese nula (H_0) onde a

comparação desejada deve ser verificada e, uma hipótese alternativa (H_1) distinta da hipótese H_0 . As possíveis decisões a serem tomadas podem levar em uma das possibilidades:

- i) H_0 é verdadeira e é aceita;
- ii) H_0 é verdadeira e é rejeitada;
- iii) H_0 é falsa e é rejeitada;
- iv) H_0 é falsa e é aceita.

Se a decisão for a (ii), diz-se que foi cometido um erro tipo I. A probabilidade de se cometer um erro tipo I é denotado por α e chamado nível de significância do teste. Se a decisão for a (iv), diz-se que foi cometido um erro tipo II. A probabilidade para o erro tipo II é denotada por β e $(1-\beta)$ é chamado potência do teste. Deve haver um balanceamento nos valores de α e β , porque diminuindo um deles o outro aumenta e vice-versa. Geralmente, na prática, utiliza-se os valores $\alpha = 5\%$, 2% ou 1% .

A hipótese mais ampla sobre as observações GPS (ou geodésicas de modo geral) é que elas são realizações amostrais de variáveis aleatórias normalmente distribuídas. No caso de uma variável aleatória (\mathbf{X}) n -dimensional, a função densidade de probabilidade normal multivariada (ou multinormal) associada é:

$$f(\mathbf{x}) = \left[\frac{1}{(2\pi)^{n/2} \sqrt{|\Sigma_x|}} \right] e^{\left[-\frac{1}{2}(\mathbf{x}-\mu_x)^T \Sigma_x^{-1} (\mathbf{x}-\mu_x) \right]} \quad (3.191)$$

onde:

Σ_x : é a matriz variância-covariância;

$\mu_x = E\{\mathbf{X}\}$.

A notação empregada para variáveis aleatórias normais multivariadas \mathbf{X} com média μ_x e matriz variância-covariância Σ_x é:

$$\mathbf{X} \sim N(\mu_x, \Sigma_x) \quad (3.192)$$

Para o modelo $V=Ax-L$ supõe-se a priori que a variável aleatória (V), tem distribuição normal multivariada, com média $E\{V\}=0$ e matriz variância-covariância $\Sigma_V = \Sigma_L$

$$\begin{aligned} V &\sim N(0, \Sigma_L) \\ &\sim N(0, \sigma_0^2 Q_L). \end{aligned} \quad (3.193)$$

Com essas suposições demonstra-se que a forma quadrática $\hat{V}^T Q_L^{-1} \hat{V} / \sigma_0^2$ tem distribuição qui-quadrado²¹, χ_v^2 , onde $v=(m-n)$ representa graus de liberdade. (Prova: ver LEICK, 1995, p. 137-142).

As hipóteses mais empregadas com base na distribuição χ_v^2 de $\hat{V}^T Q_L^{-1} \hat{V} / \sigma_0^2$, são:

$$\begin{aligned} H_0 : \sigma_0^2 &= \hat{\sigma}_0^2 \\ H_1 : \begin{cases} a) \sigma_0^2 \neq \hat{\sigma}_0^2 \\ b) \hat{\sigma}_0^2 > \sigma_0^2 \end{cases} \end{aligned}$$

A estatística do teste é:

$$\chi_*^2 = \frac{\hat{\sigma}_0^2}{\sigma_0^2} (m-n) = \frac{\hat{V}^T Q_L^{-1} \hat{V}}{\sigma_0^2}.$$

Conclusão:

- a) se $\chi_*^2 < \chi_{v, 1-\frac{\alpha}{2}}^2$ ou $\chi_*^2 > \chi_{v, \frac{\alpha}{2}}^2$ a hipótese nula H_0 é rejeitada.
- b) Se $\chi_*^2 > \chi_{v, \alpha}^2$ a hipótese nula H_0 é rejeitada.

Obs.: os valores $\chi_{v, \frac{\alpha}{2}}^2$ e $\chi_{v, 1-\frac{\alpha}{2}}^2$ podem ser obtidos por integração numérica da distribuição qui-quadrado ou podem ser retirados de tabelas disponíveis na literatura.

²¹ A distribuição qui-quadrado χ_v^2 , assim como, as distribuições t-Student e a distribuição F (Fisher), são clássicas na literatura e podem ser encontradas nas referências citadas na seção 3.2 e respectivas subseções.

Existem vários outros testes estatísticos de hipóteses como, por exemplo, verificar a igualdade entre duas variâncias de referências estimadas, ou verificar a igualdade entre dois vetores de parâmetros estimados, etc. O princípio fundamental geralmente é o mesmo, ou seja, parte-se da suposição de que as variáveis aleatórias (observações) têm distribuição normal multivariada. No capítulo 4, oportunamente, serão introduzidos outros testes para validação das ambigüidades GPS.

4. RESOLUÇÃO DAS AMBIGÜIDADES GPS

Desde o início da década de 80 até os dias de hoje, muitos métodos de resolução das ambigüidades foram e vêm sendo desenvolvidos. Não se pretende nesse trabalho apresentar um histórico sobre cada um deles, nem seria possível. No entanto, uma vasta quantidade de referências, 243 para ser mais exato, foram coletadas e listadas no endereço eletrônico: <http://www.geo.tudelft.nl/mgp/people/paul/amb.html>, mantido pelo Dr. Paul de Jonge, atualizado até outubro de 1997¹.

Dentre os métodos de resolução das ambigüidades GPS, são de interesse para esse trabalho, aqueles que utilizam observações da fase da portadora em apenas uma frequência, cujas aplicações são destinadas às linhas de base curta.

Alguns destes métodos serão descritos nas próximas subseções, mantendo um certo grau de detalhamento, e enquadrados de acordo com uma das seguintes classes: métodos de inicialização, métodos de busca no domínio das coordenadas, métodos de busca no domínio das ambigüidades e métodos integrados.

4.1 MÉTODOS DE INICIALIZAÇÃO

Dentre os possíveis métodos de inicialização, destaca-se pela simplicidade e popularidade o da troca (ou intercâmbio) de antenas e, o da linha de base conhecida. Esses métodos foram muito utilizados em levantamentos cinemáticos até meados da década de 90 (ainda hoje são, mas com menor frequência), quando então, começaram a ser mais empregados os métodos de resolução rápida das ambigüidades, tais como: Fast Ambiguity Resolution Approach (FARA) e Least-squares AMBiguity Decorrelation Adjustment (LAMBDA), entre outros.

¹ Consulta realizada em: 15 ago. 2001.

4.1.1 Troca ou Intercâmbio de Antenas

O método da troca ou intercâmbio de antenas foi introduzido por REMONDI (1985/86). Para a sua realização utiliza-se uma estação de coordenadas conhecidas (A) e uma outra estação auxiliar (B), próxima (até uns 20 m). Rastreia-se por algumas épocas os sinais dos satélites (fase da portadora) com as antenas do receptor A e do receptor B nas estações A e B, respectivamente. Depois, sem perder a sintonia dos sinais, troca-se as antenas, o receptor A passa a rastrear informações (fase da portadora) na estação B e o receptor B passa a rastrear informações na estação A. Rastreando-se algumas épocas nesta posição, pode-se retornar as antenas nas posições originais. Com esse procedimento as coordenadas da estação auxiliar podem ser determinadas e, conseqüentemente, as ambigüidades podem ser resolvidas usando o método da linha de base conhecida (seção 4.1.2). A descrição matemática do processo é desenvolvida a seguir.

Utilizando-se as equações da dupla diferença de fase (equação 2.93) para uma época t_1 (antes da troca das antenas), e para uma época t_2 (depois da troca das antenas) e desconsiderando as refrações troposférica e ionosférica (devido a proximidade das estações), tem-se:

$$\Phi_{AB}^{SP}(t_1) = \rho_B^P(t_1) - \rho_A^P(t_1) - \rho_B^S(t_1) + \rho_A^S(t_1) - \lambda N_{AB}^{SP} - \varepsilon_{AB}^{SP}(t_1); \quad (4.1)$$

$$\Phi_{AB}^{SP}(t_2) = -\rho_B^P(t_2) + \rho_A^P(t_2) + \rho_B^S(t_2) - \rho_A^S(t_2) - \lambda N_{AB}^{SP} - \varepsilon_{AB}^{SP}(t_2). \quad (4.2)$$

Nas equações (4.1) e (4.2), as distâncias receptor/satélite na época t_2 têm o sinal trocado em relação à época t_1 . A dupla diferença de ambigüidades, desde que não ocorra perda de sinais, é a mesma para as duas épocas.

Subtraindo a equação (4.1) da equação (4.2), resulta:

$$\Phi_{AB}^{SP}(t_2 - t_1) = -\rho_B^P(t_2) - \rho_B^P(t_1) + \rho_A^P(t_2) + \rho_A^P(t_1) + \rho_B^S(t_2) + \rho_B^S(t_1) - \rho_A^S(t_2) - \rho_A^S(t_1) - \varepsilon. \quad (4.3)$$

As únicas incógnitas na equação (4.3) são as coordenadas da estação auxiliar

(B). Se o resíduo ε for negligenciável e forem rastreados os mesmos 4 satélites nas épocas t_1 (antes da troca) e t_2 (depois da troca), forma-se um sistema de 3 equações a 3 incógnitas, sendo possível uma solução para as coordenadas da estação auxiliar (B). Mais épocas e mais satélites observados, permitem por meio de um ajustamento, melhorar a solução.

LEICK (1995, p. 383), chama a atenção para a diferença entre usar a equação (4.3) e a equação da tripla diferença de fase (equação 2.96). Na tripla diferença, as distâncias de um receptor para um mesmo satélite em duas épocas distintas se subtraem, ao passo que na troca de antenas elas se adicionam.

4.1.2 Linha de Base Conhecida

Na equação (2.93), da dupla diferença de fase, se forem conhecidas as coordenadas de duas estações (A) e (B) próximas² o bastante para poderem ser negligenciados os termos das refrações ionosférica e troposférica, só a dupla diferença de ambigüidade (N_{AB}^{SP}) ficaria como incógnita e, se o resíduo (ε_{AB}^{SP}) também for negligenciável, a dupla diferença de ambigüidade pode ser calculada por:

$$N_{AB}^{SP} = \frac{1}{\lambda} (\rho_{AB}^{SP} - \Phi_{AB}^{SP}). \quad (4.4)$$

Apesar de teoricamente $N_{AB}^{SP} (= N_B^P - N_A^P - N_B^S + N_A^S)$ ser um número inteiro, quando calculado pela equação (4.4), terá, em geral, um valor real, e para saber se o valor real encontrado pode ou não ser arredondado para seu inteiro mais próximo, uma análise prévia pode ser realizada com base nos erros envolvidos na equação (4.4).

Se forem rastreados n satélites, pode-se formar a cada época, $(n-1)$ equações independentes de duplas diferenças de ambigüidades.

Denotando-se os satélites pelos sobrescritos 1,2,...,n, o sistema de equações de duplas diferenças de ambigüidades, para uma época, pode ser representado por:

² O quanto próximas; depende das condições atmosféricas das estações.

$$\begin{cases} N_{AB}^{12} = \frac{1}{\lambda}(\rho_{AB}^{12} - \Phi_{AB}^{12}) \\ N_{AB}^{13} = \frac{1}{\lambda}(\rho_{AB}^{13} - \Phi_{AB}^{13}) \\ \vdots \\ N_{AB}^{1n} = \frac{1}{\lambda}(\rho_{AB}^{1n} - \Phi_{AB}^{1n}) \end{cases} \quad (4.5)$$

Supondo-se que erros no lado direito do sistema (4.5) estejam nas coordenadas da estação B (a estação B pode ser a estação auxiliar usada na troca de antenas), nas coordenadas dos satélites e nos ruídos da fase da portadora. A matriz Jacobiano (**D**), pode ser arranjada da seguinte forma:

$$D = \begin{bmatrix} \frac{\partial N_{AB}^{12}}{\partial x_B} & \frac{\partial N_{AB}^{12}}{\partial y_B} & \frac{\partial N_{AB}^{12}}{\partial z_B} & \frac{\partial N_{AB}^{12}}{\partial x^2} & \frac{\partial N_{AB}^{12}}{\partial y^2} & \frac{\partial N_{AB}^{12}}{\partial z^2} & \frac{\partial N_{AB}^{12}}{\partial x^1} & \frac{\partial N_{AB}^{12}}{\partial y^1} & \frac{\partial N_{AB}^{12}}{\partial z^1} & 0 & 0 & 0 \\ \frac{\partial N_{AB}^{13}}{\partial x_B} & \frac{\partial N_{AB}^{13}}{\partial y_B} & \frac{\partial N_{AB}^{13}}{\partial z_B} & 0 & 0 & 0 & \frac{\partial N_{AB}^{13}}{\partial x^1} & \frac{\partial N_{AB}^{13}}{\partial y^1} & \frac{\partial N_{AB}^{13}}{\partial z^1} & \frac{\partial N_{AB}^{13}}{\partial x^3} & \frac{\partial N_{AB}^{13}}{\partial y^3} & \frac{\partial N_{AB}^{13}}{\partial z^3} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \frac{\partial N_{AB}^{1n}}{\partial x_B} & \frac{\partial N_{AB}^{1n}}{\partial y_B} & \frac{\partial N_{AB}^{1n}}{\partial z_B} & 0 & 0 & 0 & \frac{\partial N_{AB}^{1n}}{\partial x^1} & \frac{\partial N_{AB}^{1n}}{\partial y^1} & \frac{\partial N_{AB}^{1n}}{\partial z^1} & 0 & 0 & 0 \end{bmatrix}$$

$$\begin{bmatrix} 0 & 0 & 0 & \frac{\partial N_{AB}^{12}}{\partial \Phi_B^2} & \frac{\partial N_{AB}^{12}}{\partial \Phi_A^2} & \frac{\partial N_{AB}^{12}}{\partial \Phi_B^1} & \frac{\partial N_{AB}^{12}}{\partial \Phi_A^1} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{\partial N_{AB}^{13}}{\partial \Phi_B^1} & \frac{\partial N_{AB}^{13}}{\partial \Phi_A^1} & \frac{\partial N_{AB}^{13}}{\partial \Phi_B^3} & \frac{\partial N_{AB}^{13}}{\partial \Phi_A^3} & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \frac{\partial N_{AB}^{1n}}{\partial x^n} & \frac{\partial N_{AB}^{1n}}{\partial y^n} & \frac{\partial N_{AB}^{1n}}{\partial z^n} & 0 & 0 & \frac{\partial N_{AB}^{1n}}{\partial \Phi_B^1} & \frac{\partial N_{AB}^{1n}}{\partial \Phi_A^1} & 0 & 0 & \frac{\partial N_{AB}^{1n}}{\partial \Phi_B^n} & \frac{\partial N_{AB}^{1n}}{\partial \Phi_A^n} \end{bmatrix} \quad (4.6)$$

A matriz variância-covariância envolvendo as coordenadas da estação (B), coordenadas dos satélites e fases da portadora, para essa análise (prévia) pode ser considerada diagonal. Os seus elementos seguindo a ordem da correspondente da matriz Jacobiano (**D**), são:

$$\Sigma_L = \text{diag}(\sigma_{x_B}^2; \sigma_{y_B}^2; \sigma_{z_B}^2; \sigma_{x^2}^2; \sigma_{y^2}^2; \sigma_{z^2}^2; \sigma_{x^1}^2; \sigma_{y^1}^2; \sigma_{z^1}^2; \sigma_{x^3}^2; \sigma_{y^3}^2; \sigma_{z^3}^2; \dots; \sigma_{x^a}^2; \sigma_{y^a}^2; \sigma_{z^a}^2; \sigma_{\Phi_B^2}^2; \sigma_{\Phi_A^2}^2; \sigma_{\Phi_B^1}^2; \sigma_{\Phi_A^1}^2; \sigma_{\Phi_B^3}^2; \sigma_{\Phi_A^3}^2; \dots; \sigma_{\Phi_B^a}^2; \sigma_{\Phi_A^a}^2). \quad (4.7)$$

Da aplicação da lei de propagação das covariâncias, obtém-se:

$$\Sigma_{N_{AB}^{ij}} = \mathbf{D} \Sigma_L \mathbf{D}^T. \quad (4.8)$$

Denotando por $\sigma_{N_{AB}^{ij}}$ os desvios padrão (obtidos das variâncias na diagonal de $\Sigma_{N_{AB}^{ij}}$) das correspondentes ambigüidades (N_{AB}^{ij}) e supondo que essas têm distribuição normal, os intervalos de confiança com probabilidades³ de 90% , 95% e 99% para os quais ($N_{AB}^{ij} - \tilde{N}_{AB}^{ij}$) estariam contidos, são:

$$P[-1,645 \sigma_{N_{AB}^{ij}} < N_{AB}^{ij} - \tilde{N}_{AB}^{ij} < +1,645 \sigma_{N_{AB}^{ij}}] = 0,90; \quad (4.9)$$

$$P[-1,960 \sigma_{N_{AB}^{ij}} < N_{AB}^{ij} - \tilde{N}_{AB}^{ij} < +1,960 \sigma_{N_{AB}^{ij}}] = 0,95; \quad (4.10)$$

$$P[-2,576 \sigma_{N_{AB}^{ij}} < N_{AB}^{ij} - \tilde{N}_{AB}^{ij} < +2,576 \sigma_{N_{AB}^{ij}}] = 0,99. \quad (4.11)$$

Para que as duplas diferenças de ambigüidades (N_{AB}^{ij}) possam ser arredondadas para o seu valor inteiro mais próximo (\tilde{N}_{AB}^{ij}), de modo a ficar assegurado que somente este valor inteiro esteja contido no intervalo de confiança, por exemplo, no intervalo dado por (4.11). O desvio padrão ($\sigma_{N_{AB}^{ij}}$) deve ser pequeno, em torno de 0,2 ciclos.

Exemplo 1:

Suponha-se que o valor da dupla diferença de ambigüidades (N_{AB}^{ij}) calculado

³ Os intervalos de confiança (4.9) a (4.11), só representam as probabilidades indicadas se as variáveis aleatórias envolvidas não forem correlacionadas. Este não é o caso, deve ficar esclarecido que os intervalos só são aproximadamente relacionados com as probabilidades indicadas. Detalhes a respeito dessa discussão são encontrados, por exemplo, em: JOHNSON e WICHERN (1988, p. 187)

pela equação (4.4), seja: $N_{AB}^{ij} = 500,68$ ciclos. Usando o intervalo de confiança com 99% de probabilidade, tem-se:

$$-2,576\sigma_{N_{AB}^{ij}} < 500,68 - \tilde{N}_{AB}^{ij} < +2,576\sigma_{N_{AB}^{ij}} . \quad (4.12)$$

Para $\sigma_{N_{AB}^{ij}} = 0,3$ ciclos, haveriam dois candidatos inteiros para \tilde{N}_{AB}^{ij} que estariam contidos no intervalo de confiança. E, portanto, nenhum deles poderia ser adotado como solução. Por outro lado, com o desvio padrão $\sigma_{N_{AB}^{ij}} = 0,2$ ciclos, somente o candidato $\tilde{N}_{AB}^{ij} = 501$, estaria contido no intervalo de confiança e poderia ser adotado como solução da ambigüidade correspondente.

Exemplo 2:

A tab. (4.1) apresenta as coordenadas no WGS-84 de quatro satélites rastreados às 12^h 28^m 59,977^s (TU) do dia 10 de novembro de 1999. Também apresenta as coordenadas de duas estações denominadas (P003) e (Pc), onde foram realizados os rastreios.

TABELA 4.1 – COORDENADAS DOS SATÉLITES E DAS ESTAÇÕES PARA UMA PRÉ-ANÁLISE DA TÉCNICA DA LINHA DE BASE CONHECIDA

Ident. Satélite (PRN)	X (m)	Y (m)	Z (m)
30	2 064 149,558	-15 547 503,049	-21 611 014,187
24	18 376 061,572	-6 145 587,396	-17 850 767,358
6	-2 608 131,028	-23 476 208,900	-11 740 593,059
5	17 087 656,614	-12 062 101,342	-16 324 607,434
Ident. Estação			
P003	3 763 110,693	-4 365 209,792	-2 725 100,117
Pc	3 762 986,693	-4 365 344,086	-2 725 070,605

Os desvios padrão das coordenadas dos satélites serão considerados iguais e de acordo com a tab. (2.2) assumirão os valores $\sigma_{SAT} = 3$ m. Os desvios padrão das coordenadas da estação Pc assumirão os valores $\sigma_{Pc} = 0,08$ m. A estação P003 será considerada isenta de erros. Os desvios padrão das fases da portadora de acordo com a tab. (1.1) e tab. (2.6) assumirão os valores $\sigma_{\phi} = 3,163$ mm.

A matriz Σ_L da equação (4.7), utilizando os desvios padrão assumidos anteriormente, torna-se:

$$\Sigma_L = \text{diag}(0,0064; 0,0064; 0,0064; 9; 9; 9; 9; 9; 9; 9; 9; 9; 9; 9; 9; 9; 0,00001; 0,00001; 0,00001; 0,00001; 0,00001; 0,00001; 0,00001; 0,00001; 0,00001).$$

Com os dados da tab. (4.1) calcula-se as derivadas parciais da matriz $\mathbf{D}_{3 \times 23}$. O resultado encontrado para $\Sigma_{N_{AB}^{ij}}$ do produto $(\mathbf{D} \Sigma_L \mathbf{D}^T)$ com tais dados, foi:

$$\Sigma_{N_{AB}^{ij}} = \begin{bmatrix} 0,103099 & -0,010516 & 0,123099 \\ -0,010516 & 0,015807 & -0,034838 \\ 0,123099 & -0,034838 & 0,209388 \end{bmatrix}.$$

Os desvios padrão a partir da diagonal de $\Sigma_{N_{AB}^{ij}}$ são $\sigma_{N_{P3,Pc}^{5,30}} = 0,321$ ciclos, $\sigma_{N_{P3,Pc}^{5,24}} = 0,126$ ciclos, $\sigma_{N_{P3,Pc}^{5,6}} = 0,457$ ciclos. Baseando-se na análise do exemplo 1, apenas a dupla diferença de ambigüidade $N_{P3,Pc}^{5,24}$, poderia ser arredondada com segurança para o seu inteiro mais próximo (usando uma época). Para as demais, de acordo com a expressão do desvio padrão da média

$$\sigma_{\bar{y}} = \frac{\sigma_y}{\sqrt{n}}, \quad (4.13)$$

onde:

$$\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i;$$

σ_y : desvio padrão da variável aleatória Y,

seriam necessárias 3 épocas ($n=3$) para proceder o arredondamento da média com segurança. No entanto, se as coordenadas da estação B tiverem um desvio padrão de 5 cm, os desvios padrão das duplas diferenças de ambigüidades seriam $\sigma_{N_{P3,Pc}^{5,30}} = 0,2$ ciclos, $\sigma_{N_{P3,Pc}^{5,24}} = 0,07$ ciclos, $\sigma_{N_{P3,Pc}^{5,6}} = 0,28$ ciclos.

O resultado dessa análise mostrou que para proceder ao arredondamento das duplas diferenças de ambigüidades usando a método da linha de base conhecida, em uma época, é necessário que o erro da linha de base seja pelo menos inferior a 10 cm.

4.2 MÉTODOS DE BUSCA NO ESPAÇO DAS COORDENADAS

Métodos dessa classe são caracterizados por encontrar no espaço das coordenadas da estação incógnita (B), qual o vetor $\mathbf{x}_B = [x_B \ y_B \ z_B]^T$, corresponde à solução do ajustamento quando é injuncionado o vetor de ambigüidades corretas ($\tilde{\mathbf{a}}$).

Dentre os métodos que utilizam esse princípio, o mais popular, é sem dúvida, o método da função de ambigüidade.

4.2.1 Método da Função de Ambigüidade

Segundo HOFMANN-WELLENHOF; LICHTENEGGER e COLLINS (1994, p. 222,223) o princípio da função de ambigüidade foi proposto por COUNSELMAN e GOUREVITCH (1981)⁴, tendo sido também investigado por REMONDI (1984)⁵ e MADER (1990)⁶. Durante a última década, vários outros autores também apresentaram suas contribuições utilizando tal princípio.

Para o desenvolvimento das expressões apresentadas nesse trabalho, seguiu-se o desenvolvimento de HOFMANN-WELLENHOF; LICHTENEGGER e COLLINS (1994, p. 223-224).

Passando a equação da simples diferença de fase (2.90) para radianos, tem-se:

⁴ COUNSELMAN, C. C.; GOUREVITCH, S. A. Miniature interferometer terminals for earth surveying: ambiguity and multipath with the Global positioning System. *IEEE Transactions on Geoscience and Remote Sensing*, GE, v. 19, n.4, 1981, p. 244-252.

⁵ REMONDI, B. W. *Using the Global Positioning System (GPS) phase observable for relative geodesy: modeling, processing, and results*. University of Texas at Austin, Center for Space Research, 1984.

⁶ MADER, G. L. Ambiguity function techniques for GPS phase initialization and kinematic solutions. In: *Proceedings of the Second International Symposium on Precise Positioning with the Global Positioning System*. Ottawa, Canada, September 3-7, 1990. p. 1233-1247.

$$\varphi_{AB}^s = \left(\frac{2\pi}{\lambda} \right) \Phi_{AB}^s = \frac{2\pi}{\lambda} \left(\rho_{AB}^s + cdt_{AB} + d_{trop(AB)}^s - d_{ion(AB)}^s - \lambda N_{AB}^s - \varepsilon_{AB}^s \right). \quad (4.14)$$

Desprezando os termos $d_{trop(AB)}^s$, $d_{ion(AB)}^s$ e o resíduo ε_{AB}^s , a equação (4.14) pode ser reescrita na forma:

$$\varphi_{AB}^s - \frac{2\pi}{\lambda} \rho_{AB}^s = 2\pi f dt_{AB} - 2\pi N_{AB}^s. \quad (4.15)$$

Elevando ambos os lados da equação (4.15) à potência e^i (com $i = \sqrt{-1}$ a unidade imaginária), vem:

$$e^{i\left\{\varphi_{AB}^s - \frac{2\pi}{\lambda} \rho_{AB}^s\right\}} = e^{i\{2\pi f dt_{AB} - 2\pi N_{AB}^s\}}. \quad (4.16)$$

Usando a propriedade $e^{i\{a+b\}} = e^{ia} e^{ib}$ no lado direito de (4.16)

$$e^{i\left\{\varphi_{AB}^s - \frac{2\pi}{\lambda} \rho_{AB}^s\right\}} = e^{i2\pi f dt_{AB}} e^{-i\{2\pi N_{AB}^s\}}. \quad (4.17)$$

Considerando o fato de N_{AB}^s ser um número inteiro e usando a propriedade $e^{ia} = \cos \alpha + i \sin \alpha$, tem-se:

$$\begin{aligned} e^{-i\{2\pi N_{AB}^s\}} &= \cos(-2\pi N_{AB}^s) + i \sin(-2\pi N_{AB}^s) \\ &= 1 + i(0) = 1. \end{aligned} \quad (4.18)$$

Com esse resultado a (4.17), torna-se:

$$e^{i\left\{\varphi_{AB}^s - \frac{2\pi}{\lambda} \rho_{AB}^s\right\}} = e^{i2\pi f dt_{AB}}. \quad (4.19)$$

Na equação (4.19), a simples diferença de fase (φ_{AB}^s) está sendo formada envolvendo os receptores (A) e (B) e o satélite (S). Se forem consideradas as demais simples diferenças envolvendo os n satélites rastreados, por exemplo, os satélites

1, 2, ..., n, ter-se-á n equações da forma (4.19), como:

$$e^{i\left\{\varphi_{AB}^1 - \frac{2\pi}{\lambda}\rho_{AB}^1\right\}} = e^{i2\pi f \, dt_{AB}} ; \quad (4.20)$$

$$e^{i\left\{\varphi_{AB}^2 - \frac{2\pi}{\lambda}\rho_{AB}^2\right\}} = e^{i2\pi f \, dt_{AB}} ; \quad (4.21)$$

$$\vdots$$

$$e^{i\left\{\varphi_{AB}^n - \frac{2\pi}{\lambda}\rho_{AB}^n\right\}} = e^{i2\pi f \, dt_{AB}} . \quad (4.22)$$

Somando cada uma das equações de (4.20) à (4.22), resulta:

$$\sum_{j=1}^n e^{i\left\{\varphi_{AB}^j - \frac{2\pi}{\lambda}\rho_{AB}^j\right\}} = n \, e^{i2\pi f \, dt_{AB}} . \quad (4.23)$$

Extraindo o módulo de ambos os lados da equação (4.23), tem-se:

$$\left\| \sum_{j=1}^n e^{i\left\{\varphi_{AB}^j - \frac{2\pi}{\lambda}\rho_{AB}^j\right\}} \right\| = n . \quad (4.24)$$

Para a aplicação da expressão (4.24) considera-se a estação (A) isenta de erros. A simples diferença de fase (φ_{AB}^j em radianos) possui as flutuações probabilísticas inerentes ao fato de ser observação. As coordenadas dos satélites ($j = 1, 2, \dots, n$) e as coordenadas da estação (B) possuem seus respectivos erros. Desta forma, na prática, o módulo da soma no lado esquerdo não deve resultar igual a n. Como as incógnitas são as coordenadas da estação (B), pode-se adotá-las como sendo a solução do seguinte problema de maximização:

$$\max_{\mathbf{x}_B \in R^3} f_{\text{amb}} = \left\| \sum_{j=1}^n e^{i\left\{\varphi_{AB}^j - \frac{2\pi}{\lambda}\rho_{AB}^j\right\}} \right\| , \quad (4.25)$$

onde:

$\mathbf{x}_B = (x_B, y_B, z_B)$: coordenadas da estação incógnita B;

$\rho_{AB}^j = \rho_B^j - \rho_A^j$.

A função definida em (4.25) denominada por função de ambigüidade, pode ser

representada em sua forma equivalente como:

$$f_{\text{amb}} = \sqrt{\left[\sum_{j=1}^n \cos\left(\varphi_{AB}^j - \frac{2\pi}{\lambda} \rho_{AB}^j\right) \right]^2 + \left[\sum_{j=1}^n \text{sen}\left(\varphi_{AB}^j - \frac{2\pi}{\lambda} \rho_{AB}^j\right) \right]^2}. \quad (4.26)$$

Quando mais épocas são levadas em conta, por exemplo, n_t épocas e os mesmos n satélites forem rastreados em cada época, a função de ambigüidade pode ser colocada na forma:

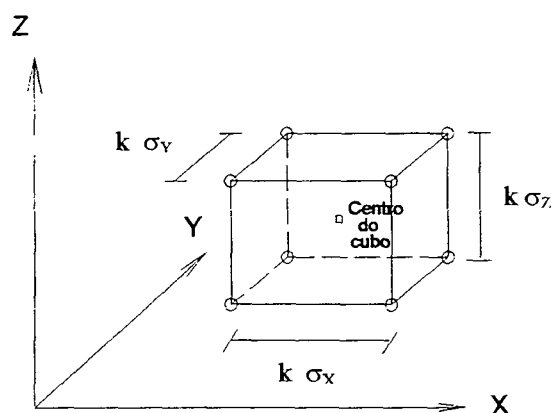
$$\max_{\mathbf{x}_B \in \mathbf{R}^3} f_{\text{amb}} = \sum_{t=1}^{n_t} \left\{ \left[\sum_{j=1}^n \cos\left(\varphi_{AB}^j - \frac{2\pi}{\lambda} \rho_{AB}^j\right) \right]^2 + \left[\sum_{j=1}^n \text{sen}\left(\varphi_{AB}^j - \frac{2\pi}{\lambda} \rho_{AB}^j\right) \right]^2 \right\}^{1/2}. \quad (4.27)$$

No caso do problema (4.27), o valor máximo da função de ambigüidade será teoricamente alcançado no valor $n.n_t$. Na prática, tal valor não deverá ser encontrado, pelos mesmos motivos já esclarecidos em (4.25).

Para resolver o problema (4.25) ou (4.27), deve-se procurar no espaço das coordenadas (X,Y,Z) , qual o conjunto de coordenadas do ponto incógnito $B(x_B, y_B, z_B)$ que maximiza a função de ambigüidade (f_{amb}).

Ocorre, porém, que a função de ambigüidade não é uma função convexa, admitindo uma infinidade de pontos de máximo sobre o \mathbf{R}^3 . É preciso então definir um conjunto $\Omega \in \mathbf{R}^3$ tal que $\mathbf{x}_B^* \in \Omega$, seja maximizador de f_{amb} . Esse conjunto pode ser definido para ser um cubo ou um elipsóide, cujo centro, pode estar na solução flutuante do ajustamento com a dupla diferença, ou então, na solução do ajustamento com a tripla diferença, ou ainda, na solução dada pelo ajustamento das pseudodistâncias (código). O tamanho do cubo (ou do elipsóide), geralmente é definido com base nas informações estatísticas (desvio padrão σ) das coordenadas ajustadas (fig. 4.1).

FIGURA 4.1 – ESPAÇO DE BUSCA DAS COORDENADAS



A constante k na fig. (4.1) é definida visando a obtenção de um intervalo de confiança desejado.

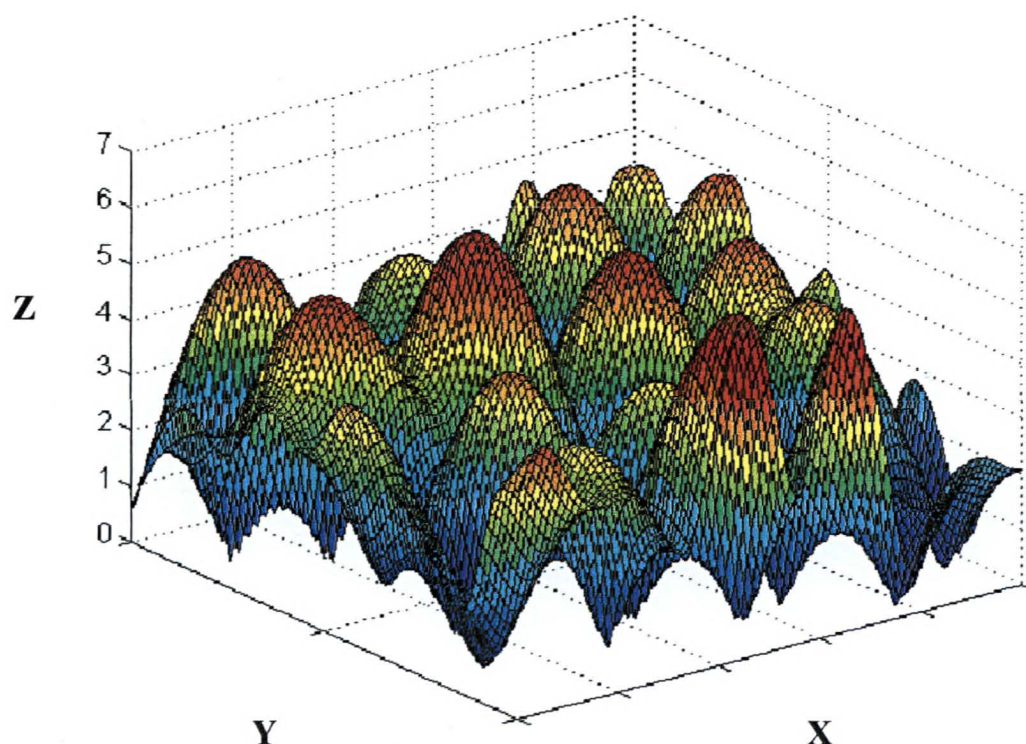
Por causa da função (f_{amb}) ser composta da soma de várias funções harmônicas, ela se torna uma função cujo gráfico é cheio de picos, dificultando a aplicação dos métodos de otimização estudados no capítulo 3. A não ser em casos particulares, como será visto na seção (4.4.2), o que se faz é formar grides regularmente espaçados testando-se cada um dos pontos do gride até encontrar o maximizador da função (f_{amb}). Uma vez encontrado este maximizador, forma-se um novo espaço centrado nele, com grides menores, e testa-se novamente os pontos desse novo gride. O valor x_B^* maximizador da função (f_{amb}) é tomado como solução. Daí, as ambigüidades podem ser calculadas pelo método da linha de base conhecida.

Para esclarecer sobre o comportamento da função de ambigüidade, será mostrado o seguinte exemplo utilizando a simples diferença de fase, 1 época, e 7 satélites rastreados.

As coordenadas corretas da estação incógnita B, são: $x_B = 3\,762\,986,693$ m, $y_B = -4\,365\,344,086$ m e $z_B = -2\,725\,070,605$ m. Fixando a coordenada $z_B = -2\,725\,070,605$ m e fazendo uma variação da coordenada x_B de 0,01 a partir de

3 762 986 m até 3 762 987 m e uma variação da coordenada y_B de 0,01 a partir de -4 365 344 m até -4 365 343 m, resulta no gráfico da fig. (4.2).

FIGURA 4.2 – GRÁFICO DA FUNÇÃO DE AMBIGÜIDADE



Nota-se na fig. (4.2), uma grande quantidade de picos a serem investigados em apenas 1 m^2 no plano (X,Y) . Esse exemplo mostra claramente que quanto maior o espaço de busca, muito mais lenta se torna a procura.

Se os métodos de otimização (ver seção 3.1.7.4, lembrando que $\max f_{\text{amb}} = \min - f_{\text{amb}}$) fossem aplicados, iriam encontrar os máximos locais a partir do ponto inicial dado. Desta forma, nenhuma garantia sobre o máximo global poderia ser assegurado para a função de ambigüidade, a não ser, que a busca fosse realizada em todo o domínio $(X,Y) \in \mathbb{R}^2$.

Outra característica da função de ambigüidade é que ela é insensível em relação à perda de ciclos, já que $\cos(k\varphi) = \cos(\varphi)$ e $\sin(k\varphi) = \sin(\varphi)$, para todo $k \in \mathbb{Z}$.

Cabe observar que as equações desenvolvidas nessa seção utilizam apenas a

simples diferença de fase, mas, nada impede em se utilizar a dupla diferença de fase, com faz LEICK (1995, p. 375), CORBETT e CROSS (1995, p. 152).

4.3 MÉTODOS DE BUSCA NO ESPAÇO DAS AMBIGÜIDADES

Nesse tipo de método procura-se resolver as ambigüidades trabalhando-se diretamente sobre elas, e não se utiliza, por exemplo, artifícios de busca das coordenadas como foi nos métodos da linha de base conhecida e da função de ambigüidade.

Os métodos de busca sobre espaço das ambigüidades, geralmente, aceitam como solução, após uma validação estatística, a solução do problema de mínimos quadrados inteiro. Algumas destas técnicas são descritas nas próximas subseções.

4.3.1 Solução Baseada no Desvio Padrão das Ambigüidades.

O método mais simples é, após a obtenção da solução flutuante, arredondar as ambigüidades para seu valor inteiro mais próximo. No entanto, esse tipo de estratégia só pode ser aplicado com sucesso para sessões de longa duração (de 30 minutos a várias horas) e para linhas de base curta. Mas, mesmo assim, se nenhuma outra informação puder confirmar os resultados, ficariam dúvidas sobre a correta resolução das ambigüidades.

Por outro lado, um tipo de análise semelhante aquela dos exemplos 1 e 2 da seção 4.1.2, poderia ser aplicada com as informações provenientes do ajustamento: ambigüidades ajustadas da solução flutuante ($\hat{\mathbf{x}}_2$), matriz variância-covariância estimada ($\Sigma_{\hat{\mathbf{x}}_2}$), graus de liberdade do ajustamento ($v=m-n$) e a variância de referência a posteriori ($\hat{\sigma}_0^2$).

Considere N_i e N_j duas duplas diferenças de ambigüidades contidas no vetor \mathbf{x}_2 . Os desvios padrão estimados para N_i e para a diferença $N_{ij}=N_i-N_j$ são, respectivamente:

$$\hat{\sigma}_{N_i} = \sqrt{\Sigma_{\hat{\mathbf{x}}_2(i,i)}}; \quad (4.28)$$

$$\hat{\sigma}_{N_{ij}} = \sqrt{\sum_{i_{2(i,j)}} - 2 \sum_{i_{2(i,j)}} + \sum_{i_{2(i,j)}}}, \quad (4.29)$$

onde:

$\sum_{i_{2(i,i)}}$: variância da ambigüidade N_i ;

$\sum_{i_{2(i,j)}}$: covariância entre as duplas diferenças de ambigüidades N_i e N_j .

As ambigüidades N_i e N_j poderiam ser arredondadas para seus valores inteiros mais próximos, se:

- i) a hipótese $H_0: \sigma_0^2 = \hat{\sigma}_0^2$, não fosse rejeitada pelo teste qui-quadrado (ver seção 3.2.6) a um nível de significância α (por exemplo $\alpha = 5\%$); e
- ii) os intervalos de confiança $P[\hat{N}_i - t_{v, 1-\frac{\alpha}{2}} \hat{\sigma}_{N_i} \leq N_i \leq \hat{N}_i + t_{v, \frac{\alpha}{2}} \hat{\sigma}_{N_i}] = 1 - \alpha$ e $P[\hat{N}_{ij} - t_{v, 1-\frac{\alpha}{2}} \hat{\sigma}_{N_{ij}} \leq N_{ij} \leq \hat{N}_{ij} + t_{v, \frac{\alpha}{2}} \hat{\sigma}_{N_{ij}}] = 1 - \alpha$, contiverem apenas um número inteiro,

onde:

$t_{v, 1-\frac{\alpha}{2}}, t_{v, \frac{\alpha}{2}}$: são os limites teóricos (percentis) inferior e superior, com base na distribuição t-Student com v graus de liberdade e nível de significância α .

Cumpridas as condições anteriores (i) e (ii), as ambigüidades \hat{N}_i e \hat{N}_j podem ser fixadas para o seus valores inteiros mais próximos (\tilde{N}_i) e (\tilde{N}_j), e um novo ajustamento pode ser realizado com essas ambigüidades consideradas como quantidades conhecidas (injuncionadas). Esse procedimento ajudaria na resolução das demais ambigüidades e deveria ser aplicado até que todas as ambigüidades tenham sido resolvidas, ou, se nenhuma das ambigüidades fosse resolvida por esse processo, seriam necessárias observações de uma nova época.

A técnica descrita nessa seção está implementada, por exemplo, no *software* Bernese V.4.0 (ver ROTHACHER et al., 1996, p. 203).

4.3.2 Fast Ambiguity Resolution Approach (FARA)

O método FARA, devido a FREI e BEUTLER (1990); FREI (1991, p. 63), foi um dos primeiros a ser popularizado para aplicações no modo estático rápido. Fica caracterizado pelas 3 fases (i), (ii) e (iii) a seguir:

- i) Cada uma das ambigüidades da dupla diferença N_i , $i = 1, 2, \dots, t$ contidas no vetor t -dimensional \mathbf{x}_2 , é definida pelo intervalo de confiança unidimensional:

$$P[\hat{N}_i - t_{v, 1-\frac{\alpha}{2}} \hat{\sigma}_{N_i} \leq N_i \leq \hat{N}_i + t_{v, \frac{\alpha}{2}} \hat{\sigma}_{N_i}] = 1-\alpha, \quad (4.30)$$

onde:

\hat{N}_i : valor da i -ésima ambigüidade estimada na solução flutuante;

$v = m-n$: graus de liberdade do ajustamento;

$t_{v, 1-\frac{\alpha}{2}}$, $t_{v, \frac{\alpha}{2}}$: percentis inferior e superior, com base na distribuição

t -Student com v graus de liberdade e nível de significância α ;

$N_{ij} = N_i - N_j$: diferença das ambigüidades.

Por exemplo: se no intervalo (4.30) houver 4 valores inteiros para N_i , então, os 4 valores serão candidatos a solução e deverão ser testados.

- ii) As ambigüidades inteiras (candidatas à solução) são ordenadas em grupo conforme ilustra o exemplo a seguir.

Suponha-se que o vetor das ambigüidades resultantes da solução flutuante, seja: $\hat{\mathbf{x}}_2 = [4,9 \ 18,1 \ 30,7]^T$ e, que estão contemplados pelo intervalo de

confiança (4.30) os seguintes valores para cada uma delas:

$$N_1 = \{4 \ 5 \ 6\}; \quad N_2 = \{17 \ 18 \ 19\}; \quad N_3 = \{29 \ 30 \ 31 \ 32 \ 33\}.$$

Os grupos com cada vetor de ambigüidades inteiras a serem testados são (ver também LEICK, 1995, p. 359):

5 18 31	4 18 31	6 18 31
5 18 30	4 18 30	6 18 30
5 18 32	4 18 32	6 18 32
5 18 29	4 18 29	6 18 29
5 18 33	4 18 33	6 18 33
5 19 31	4 19 31	6 19 31
5 19 30	4 19 30	6 19 30
5 19 32	4 19 32	6 19 32
5 19 29	4 19 29	6 19 29
5 19 33	4 19 33	6 19 33
5 17 31	4 17 31	6 17 31
5 17 30	4 17 30	6 17 30
5 17 32	4 17 32	6 17 32
5 17 29	4 17 29	6 17 29
5 17 33	4 17 33	6 17 33

Observa-se que na ordenação do exemplo, a ambigüidade N_3 varia mais rapidamente, depois é seguida pela ambigüidade N_2 e, finalmente pela ambigüidade N_1 .

O número total de vetores de ambigüidades a serem testados é (FREI; BEUTLER, 1990, p. 339):

$$N_{(1)} = \prod_{i=1}^t (n_i), \quad (4.31)$$

onde:

n_i = número de valores inteiros para a ambigüidade N_i (para o exemplo, $n_1 = 3$, $n_2 = 3$ e $n_3 = 5$);
 t = número de ambigüidades (componentes do vetor \mathbf{x}_2).

Testa-se a diferença de ambigüidades ($N_{12} = N_1 - N_2$) para o intervalo de confiança:

$$P\left[\hat{N}_{12} - t_{v, 1-\frac{\alpha}{2}} \hat{\sigma}_{N12} \leq N_{12} \leq \hat{N}_{12} + t_{v, \frac{\alpha}{2}} \hat{\sigma}_{N12}\right] = 1 - \alpha. \quad (4.32)$$

Se a diferença N_{12} não pertencer ao intervalo. Descarta-se todos os vetores subseqüentes da lista que possuem os valores N_1 e N_2 para as ambigüidades, num total de:

$$N_{(2)} = \prod_{i=3}^t (n_i). \quad (4.33)$$

Se a diferença N_{12} pertencer ao intervalo (4.32), faz-se à diferença entre a primeira e a terceira ambigüidade, e verifica-se com o intervalo correspondente. Se as diferenças não pertencerem ao intervalo, descartam-se os vetores que possuem como candidatos tais valores.

O algoritmo procede desta forma até que todos os grupos de ambigüidades sejam verificados. O resultado final do procedimento é uma lista de vetores de ambigüidades inteiras, candidatos potencialmente à solução.

- iii) Os vetores candidatos potencialmente à solução, são introduzidos novamente no ajustamento como quantidades conhecidas e aquele vetor que fornecer a menor variância de referência a posteriori:

$$\tilde{\sigma}_0^2 = \frac{\tilde{\mathbf{V}}^T \mathbf{Q}_L^{-1} \tilde{\mathbf{V}}}{m - n - t}, \quad (4.34)$$

é a solução, a menos que as hipóteses nulas:

$H_0 : \tilde{\sigma}_0^2 = \sigma_0^2$ for rejeitada pelo teste estatístico qui-quadrado (ver seção

3.2.6) a um nível de significância, por exemplo, $\alpha=5\%$; e

$H_0 : \tilde{\sigma}_0^2 = \tilde{\sigma}_0^{2'}$ não for rejeitada. H_0 é rejeitada a um nível de significância

α , quando: $\frac{\tilde{\sigma}_0^2}{\tilde{\sigma}_0^{2'}} < F_{v_1, v_2, 1-\alpha}$, com $v_1=v_2=v=m-n-t$; e $F_{v_1, v_2, 1-\alpha}$ é

o percentil da distribuição F-Fisher.

Na última hipótese nula, $\tilde{\sigma}_0^{2'}$, é a segunda menor variância de referência a

posteriori encontrada no ajustamento com as ambigüidades injuncionadas.

O método FARA, aqui descrito resumidamente, está implementado no *software* Bernese V. 4.0 (ver ROTHACHER et al. 1996, p. 204).

4.3.3 Busca com a Decomposição de Cholesky

Na terceira fase do método FARA, calcula-se o vetor de resíduos com as ambigüidades fixadas e forma-se a quadrática:

$$\Omega = \check{\mathbf{V}}^T \mathbf{Q}_L^{-1} \check{\mathbf{V}}. \quad (4.35)$$

O vetor dos resíduos ($\check{\mathbf{V}}$) é calculado por :

$$\check{\mathbf{V}} = \mathbf{A} \check{\mathbf{x}} - \mathbf{L}, \quad (4.36)$$

e, de acordo com as equações (3.90), (3.91) e (3.94), escreve-se:

$$\check{\mathbf{x}} = \hat{\mathbf{x}} + \Delta \mathbf{x}, \quad (4.37)$$

sendo:

$$\hat{\mathbf{x}} = \mathbf{Q}_{\hat{\mathbf{x}}} \mathbf{A}^T \mathbf{Q}_L^{-1} \mathbf{L} \text{ (solução flutuante);} \quad (4.38)$$

$$\Delta \mathbf{x} = \mathbf{Q}_{\hat{\mathbf{x}}} \mathbf{G}^T [\mathbf{G} \mathbf{Q}_{\hat{\mathbf{x}}} \mathbf{G}^T]^{-1} (\mathbf{a} - \mathbf{G} \hat{\mathbf{x}}) \text{ (influência das injunções (a)).} \quad (4.39)$$

Substituindo a equação (4.37) na (4.36), e posteriormente levando em (4.35), e considerando $\hat{\mathbf{V}} = \mathbf{A} \hat{\mathbf{x}} - \mathbf{L}$, a forma quadrática Ω , pode ser reescrita como:

$$\Omega = \hat{\mathbf{V}}^T \mathbf{Q}_L^{-1} \hat{\mathbf{V}} + 2 \hat{\mathbf{V}}^T \mathbf{Q}_L^{-1} \mathbf{A} \Delta \mathbf{x} + \Delta \mathbf{x}^T \mathbf{A}^T \mathbf{Q}_L^{-1} \mathbf{A} \Delta \mathbf{x}. \quad (4.40)$$

E vista como soma das parcelas:

$$\Omega_1 = \hat{\mathbf{V}}^T \mathbf{Q}_L^{-1} \hat{\mathbf{V}}; \text{ e} \quad (4.41)$$

$$\Omega_2 = 2 \hat{\mathbf{V}}^T \mathbf{Q}_L^{-1} \mathbf{A} \Delta \mathbf{x} + \Delta \mathbf{x}^T \mathbf{A}^T \mathbf{Q}_L^{-1} \mathbf{A} \Delta \mathbf{x}. \quad (4.42)$$

A forma quadrática (Ω_1), é calculada com o vetor dos resíduos ($\hat{\mathbf{V}} = \mathbf{A} \hat{\mathbf{x}} - \mathbf{L}$), sem a influência das injunções (\mathbf{a}). E a forma quadrática (Ω_2), corresponde a influência das injunções (\mathbf{a}) na forma quadrática (Ω).

Substituindo na equação (4.42), a equação $\hat{\mathbf{V}} = \mathbf{A}\hat{\mathbf{x}} - \mathbf{L}$, e utilizando a forma particionada de $\mathbf{G} = [\mathbf{0} \ \mathbf{I}]$, encontra-se para Ω_2 a expressão (para mais detalhes ver seção 3.1.5):

$$\Omega_2 = (\mathbf{a} - \hat{\mathbf{x}}_2)^T \mathbf{Q}_{\hat{\mathbf{x}}_2}^{-1} (\mathbf{a} - \hat{\mathbf{x}}_2), \quad (4.43)$$

onde:

$$\mathbf{Q}_{\hat{\mathbf{x}}_2}^{-1} = \mathbf{A}_2^T \mathbf{Q}^{-1} \mathbf{A}_2, \text{ com:}$$

$$\mathbf{Q}^{-1} = \mathbf{Q}_L^{-1} - \mathbf{Q}_L^{-1} \mathbf{A}_1 (\mathbf{A}_1^T \mathbf{Q}_L^{-1} \mathbf{A}_1)^{-1} \mathbf{A}_1^T \mathbf{Q}_L^{-1}.$$

A forma quadrática Ω_2 corresponde a quadrática (função objetivo) do problema de mínimos quadrados inteiro (equação 3.82).

Do exposto anteriormente, o cálculo da variância de referência a posteriori ($\tilde{\sigma}_0^2$) (equação 4.34), pode ser calculada como:

$$\tilde{\sigma}_0^2 = \frac{\Omega_1 + \Omega_2}{m - n - t}. \quad (4.44)$$

Uma vez que Ω_1 é calculada com a solução flutuante ($\hat{\mathbf{x}}$), para encontrar $\tilde{\sigma}_0^2$ (mínima), basta encontrar $\tilde{\mathbf{a}} = \tilde{\mathbf{x}}_2$ solução do problema de mínimos quadrados inteiro (3.82).

Por ser a matriz $\mathbf{Q}_{\hat{\mathbf{x}}_2}^{-1}$ simétrica e definida positiva (ver equação 3.73), ela admite a decomposição de Cholesky

$$\mathbf{E}\mathbf{E}^T = \mathbf{Q}_{\hat{\mathbf{x}}_2}^{-1}, \quad (4.45)$$

e a forma quadrática (4.43), pode ser expressa por:

$$\Omega_2 = (\mathbf{a} - \hat{\mathbf{x}}_2)^T \mathbf{E} \mathbf{E}^T (\mathbf{a} - \hat{\mathbf{x}}_2) = \|\mathbf{E}^T (\mathbf{a} - \hat{\mathbf{x}}_2)\|_2^2 = \sum_{i=1}^l (g_i)^2. \quad (4.46)$$

A decomposição de Cholesky em (4.45), só precisa ser realizada uma vez, e o valor mínimo para Ω_2 , é encontrado testando-se cada vetor de ambigüidades, candidato potencialmente a solução.

Contudo, para checar um candidato ao mínimo da forma quadrática (Ω_2), não

é necessário calcular o somatório completo $(\sum_{i=1}^t (g_i)^2)$. Pode-se descartá-lo antecipadamente, desde que, o valor da soma para tal candidato ultrapasse o limite pré-estabelecido pela desigualdade da equação (4.47) a seguir (LEICK, 1995, p. 361):

$$\frac{\sum (g_i)^2}{\Omega_1} \frac{m-n}{t} \leq F_{t,m-n,\alpha}, \quad (4.47)$$

onde:

$F_{t,m-n,\alpha}$: percentil superior com base na distribuição F-Fisher, com t e (m-n) graus de liberdade e nível de significância α ;

Ω_1 : forma quadrática da solução flutuante.

Obs.: a rigor, a estatística do lado esquerdo de (4.47) segue apenas aproximadamente a distribuição $F_{t,m-n,\alpha}$, mas ao propósito de selecionar previamente os candidatos, pode ser aplicada sem maiores prejuízos.

A forma de procura baseada na decomposição de Cholesky foi proposta por EULER e LANDAU (1992), como um método de resolução de ambigüidades *On-The-Fly* (OTF). No entanto, segundo MONICO (1996, p. 469), existe um problema na proposta inicial de procura de EULER e LANDAU (1992) para aplicações OTF, em relação ao tamanho do cubo t-dimensional. Quando, por exemplo, $2w = 50$ ciclos é a largura do cubo e $t = 6$ o número de ambigüidades, haveria $(2w+1)^t = 1.10^{12}$ vetores de ambigüidades para serem testados, o que inviabilizaria a aplicação OTF.

De acordo com LEICK (1995, p. 362), usando a estratégia de busca com a decomposição de Cholesky, e a forma de ordenação do método FARA - fase (ii), o teste em relação aos candidatos é acelerado, uma vez que, quando o teste (4.47) descarta algum candidato antecipadamente, todos os demais candidatos que possuem os mesmos valores, nas mesmas posições, podem ser descartados.

4.3.4 Least-Squares Ambiguity Decorrelation Adjustment (LAMBDA)

O método LAMBDA para a resolução das ambigüidades GPS foi inicialmente

apresentado em TEUNISSEN (1993)⁷. As implementações numéricas foram realizadas na Universidade de Delft, Holanda, sob a supervisão do prof. Teunissen por Jonge e Tiberius, apud JONGE e TIBERIUS (1996).

A principal característica do método é a solução do problema de mínimos quadrados inteiro:

$$\min_{\mathbf{a} \in \mathbb{Z}^t} \Omega_2(\mathbf{a}) = (\mathbf{a} - \hat{\mathbf{x}}_2)^T \mathbf{Q}_{\hat{\mathbf{x}}_2}^{-1} (\mathbf{a} - \hat{\mathbf{x}}_2), \quad (4.48)$$

por meio de uma reparametrização das ambigüidades, de modo que no novo espaço, se tornem menos correlacionadas do que no espaço original. Com isso, segundo TEUNISSEN; JONGE e TIBERIUS (1995, p.60) e RIZOS e HAN (1995, p. 26), a busca discreta ao minimizador inteiro ($\tilde{\mathbf{a}}$) torna-se mais eficiente.

Na seção (3.1.5), foi introduzido um exemplo mostrando que se as ambigüidades da solução flutuante ($\hat{\mathbf{a}} = \hat{\mathbf{x}}_2$) não fossem correlacionadas, então, para encontrar o vetor de ambigüidades ($\tilde{\mathbf{a}}$), minimizador de (4.48), bastaria arredondar cada um dos componentes reais de $\hat{\mathbf{a}}$ para seu valor inteiro mais próximo.

Na prática, usando observações da dupla diferença de fase em um ajustamento seqüencial, a matriz $\mathbf{Q}_{\hat{\mathbf{x}}_2}^{-1}$ é completa, indicando correlação entre as ambigüidades. Além disso, por causa do mal-condicionamento do sistema de equações normais (ver fig. 3.3), o elipsóide t-dimensional definido pela equação (4.49) é bastante alongado.

O hiper-elipsóide t-dimensional de busca é definido como:

$$(\mathbf{a} - \hat{\mathbf{x}}_2)^T \mathbf{Q}_{\hat{\mathbf{x}}_2}^{-1} (\mathbf{a} - \hat{\mathbf{x}}_2) \leq \chi^2, \quad (4.49)$$

onde:

χ^2 : é uma constante positiva escolhida convenientemente, de modo a ficar assegurado que $\tilde{\mathbf{a}}$ esteja no interior do elipsóide.

⁷ TEUNISSEN, P. J. G. Least-squares estimation of the integer GPS ambiguities. Invited lecture. Section IV – Theory and Methodology, IAG – General Meeting. Beijing, China, aug. 1993.

A reparametrização do método LAMBDA é feita mediante a transformação-Z, como segue:

$$\begin{cases} \mathbf{z} = \mathbf{Z}^T \mathbf{a} \\ \hat{\mathbf{z}} = \mathbf{Z}^T \hat{\mathbf{a}} \end{cases} \quad (4.50)$$

A matriz cofatora das ambigüidades reais transformadas (\mathbf{Q}_z), é obtida aplicando-se a lei de propagação das covariâncias (ou propagação das cofatoras). Assim,

$$\mathbf{Q}_z = \mathbf{Z}^T \mathbf{Q}_{\hat{\mathbf{z}}} \mathbf{Z}. \quad (4.51)$$

Se os elementos da matriz \mathbf{Z} forem números inteiros e, $\det(\mathbf{Z}) = \pm 1$, os elementos de sua inversa (\mathbf{Z}^{-1}) também serão números inteiros (TEUNISSEN, 1995, p. 494). Desta forma, com as equações (4.50) e (4.51) convenientemente substituídas na equação (4.49), obtém-se o novo espaço de busca:

$$(\mathbf{z} - \hat{\mathbf{z}})^T \mathbf{Q}_z^{-1} (\mathbf{z} - \hat{\mathbf{z}}) \leq \chi^2. \quad (4.52)$$

A matriz \mathbf{Z} tem que ser escolhida de alguma maneira de modo a ser mantida as propriedades anteriormente citadas e que a matriz cofatora \mathbf{Q}_z seja diagonal, ou o mais próximo possível disto, a fim diminuir a correlação entre as ambigüidades transformadas.

JONGE e TIBERIUS (1996, p. 7-17) descrevem detalhadamente um método para obter a matriz de transformação \mathbf{Z} , a partir da decomposição \mathbf{LDL}^T da matriz cofatora $\mathbf{Q}_{\hat{\mathbf{z}}}$, utilizando uma seqüência de transformações inteira de Gauss, com permutações de linhas e colunas.

Um outro método para obter a transformação-Z foi proposto por RIZOS e HAN (1995, p. 1149). Nele a matriz \mathbf{Z} é obtida iterativamente pelos produtos:

$$\mathbf{Z} = \mathbf{Z}_{\mathbf{L}_{k-1}} \mathbf{Z}_{\mathbf{U}_{k-1}} \dots \mathbf{Z}_{\mathbf{L}_1} \mathbf{Z}_{\mathbf{U}_1}, \quad (4.53)$$

onde:

$$\mathbf{Z}_{\mathbf{U}_1} = [\text{int}(\mathbf{U}_1)]^{-1},$$

com: $\mathbf{U}_1 \mathbf{D}_{\mathbf{U}_1} \mathbf{U}_1^T$ a decomposição triangular superior unitária de $\mathbf{Q}_{\hat{\mathbf{x}}_2}$;

$$\mathbf{Z}_{\mathbf{L}_1} = [\text{int}(\mathbf{L}_1)]^{-1},$$

com: $\mathbf{L}_1 \mathbf{D}_{\mathbf{L}_1} \mathbf{L}_1^T$ a decomposição triangular inferior unitária de $\mathbf{Q}_{\hat{\mathbf{z}}, \mathbf{U}_1}$

sendo: $\mathbf{Q}_{\hat{\mathbf{z}}, \mathbf{U}_1} = \mathbf{Z}_{\mathbf{U}_1} \mathbf{Q}_{\hat{\mathbf{x}}_2} \mathbf{Z}_{\mathbf{U}_1}^T$;

$$\mathbf{Q}_{\hat{\mathbf{z}}, \mathbf{L}_1} = \mathbf{Z}_{\mathbf{L}_1} \mathbf{Q}_{\hat{\mathbf{z}}, \mathbf{U}_1} \mathbf{Z}_{\mathbf{L}_1}^T.$$

O processo continua para $\mathbf{Z}_{\mathbf{U}_i}$ obtido a partir de $\mathbf{Q}_{\hat{\mathbf{z}}, \mathbf{L}_{i-1}}$ ($\mathbf{Q}_{\hat{\mathbf{x}}_2}$ é para quando $i=1$) e $\mathbf{Z}_{\mathbf{L}_i}$ obtido a partir de $\mathbf{Q}_{\hat{\mathbf{z}}, \mathbf{U}_i}$, $i=1, 2, \dots, k$, até que ambas matrizes $\text{int}(\mathbf{U}_k)$ e $\text{int}(\mathbf{L}_k)$ se tornem matrizes identidades.

Uma vez obtida a matriz de transformação \mathbf{Z} , calcula-se $\hat{\mathbf{z}}$ pela segunda das equações (4.50) e $\mathbf{Q}_{\hat{\mathbf{z}}}$ pela equação (4.51).

O valor da constante (χ^2) limitante do espaço de busca pode ser definido, por exemplo, como (ver equação 4.47):

$$\chi^2 = \frac{t}{m-n} \Omega_1 F_{t, m-n, \alpha}. \quad (4.54)$$

JONGE e TIBERIUS (1996, p. 29) descrevem duas outras estratégias para definir o valor χ^2 , relacionando o número de vetores candidatos e o volume do espaço elipsoidal. No caso, mostraram ser mais eficiente usar o espaço transformado (4.52) do que usar o espaço original (4.49).

Depois de definido o valor constante (χ^2) que define o tamanho do elipsóide t-dimensional de busca, deve-se definir os vetores de candidatos (\mathbf{z}), onde um deles, $\tilde{\mathbf{z}}$, deve ser a solução do problema transformado:

$$\min_{\mathbf{z} \in \mathbf{Z}^t} \Omega_2(\mathbf{z}) = (\mathbf{z} - \hat{\mathbf{z}})^T \mathbf{Q}_{\hat{\mathbf{z}}}^{-1} (\mathbf{z} - \hat{\mathbf{z}}) \leq \chi^2. \quad (4.55)$$

Para encontrar os vetores candidatos de ambigüidades transformadas (\mathbf{z}), JONGE e TIBERIUS (1996, p. 20-22) utilizam a decomposição \mathbf{LDL}^T de \mathbf{Q}_z^{-1} , com a quadrática (4.55) escrita na forma algébrica:

$$\sum_{i=1}^t d_i [(z_i - \hat{z}_i) + \sum_{j=i+1}^t l_{ji} (z_j - \hat{z}_j)]^2 \leq \chi^2, \quad (4.56)$$

e desenvolvem a expressão que fornece o limite para uma ambigüidade particular (z_i), sendo dado pela expressão:

$$\hat{z}_i - \sqrt{\text{right}_i} - \sum_{j=i+1}^t l_{ji} (z_j - \hat{z}_j) \leq z_i \leq \hat{z}_i + \sqrt{\text{right}_i} - \sum_{j=i+1}^t l_{ji} (z_j - \hat{z}_j), \quad (4.57)$$

onde:

$$\text{right}_i = \frac{\chi^2}{d_i} - \frac{1}{d_i} \sum_{k=i+1}^t d_k [(z_k - \hat{z}_k) + \sum_{j=k+1}^t l_{jk} (z_j - \hat{z}_j)]^2. \quad (4.58)$$

As ambigüidades (z_i), componentes dos vetores candidatos (\mathbf{z}), são encontradas usando a expressão (4.57) numa forma recursiva, isto é, a partir da última ambigüidade (z_t), até a primeira ambigüidade (z_1).

Com o procedimento anterior, todos os vetores de ambigüidades (\mathbf{z}), dentro do espaço elipsoidal (4.49) ficam determinados. O próximo passo é encontrar qual deles é o minimizador ($\tilde{\mathbf{z}}$) do problema de mínimos quadrados inteiro transformado (4.55). Para isso, deve-se testar, um, a um, todos os candidatos.

A estratégia adotada por JONGE e TIBERIUS (1996, p. 23) foi usar a decomposição \mathbf{LDL}^T de \mathbf{Q}_z^{-1} , para calcular a norma $t(\mathbf{z}) = \|\mathbf{z} - \hat{\mathbf{z}}\|_{\mathbf{Q}_z^{-1}}^2$, aproveitando a sua estrutura na forma algébrica:

$$t(\mathbf{z}) = \sum_{i=1}^t d_i [(z_i - \hat{z}_i) + \sum_{j=i+1}^t l_{ji} (z_j - \hat{z}_j)]^2. \quad (4.59)$$

Desenvolvendo a equação (4.59) para $i=1$, tem-se:

$$t(\mathbf{z}) = d_1 [(z_1 - \hat{z}_1) + \sum_{j=2}^t l_{j1} (z_j - \hat{z}_j)]^2 + \sum_{i=2}^t d_i [(z_i - \hat{z}_i) + \sum_{j=i+1}^t l_{ji} (z_j - \hat{z}_j)]^2. \quad (4.60)$$

Se um candidato qualquer (\mathbf{z}) for testado, então, para esse candidato, o valor da quadrática $t(\mathbf{z})$ é disponibilizado.

Suponha, agora, que exista um outro candidato (\mathbf{z}') para ser testado, e apenas a primeira componente de \mathbf{z} e \mathbf{z}' diferem (considerando, por exemplo, os dois candidatos $\mathbf{z} = [5 \ 18 \ 31]^T$ e $\mathbf{z}' = [4 \ 18 \ 31]^T$ usados na primeira linha da fase (ii) do método FARA, seção 4.3.2). Note que o segundo termo da equação (4.60), é o mesmo para \mathbf{z} e \mathbf{z}' e, no primeiro termo da quantidade entre colchetes, o somatório também é o mesmo, o que muda são as diferenças $(z_1 - \hat{z}_1)$ e $(z'_1 - \hat{z}_1)$. Os valores idênticos no cálculo das quadráticas $t(\mathbf{z})$ e $t(\mathbf{z}')$ não precisariam ser repetidos. Desta forma, a busca ficaria otimizada.

Uma outra estratégia é calcular $t(\mathbf{z}')$ acrescentando a influência de um incremento $\Delta t(\mathbf{z})$ em $t(\mathbf{z})$, devido à influência de um incremento Δ na primeira componente de \mathbf{z} . Exemplo:

$$\mathbf{z}' = \mathbf{z} + \Delta \mathbf{z}_1 = \begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix} + \begin{bmatrix} \Delta \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} z_1 + \Delta \\ z_2 \\ z_3 \end{bmatrix}. \quad (4.61)$$

Obs.: para os vetores \mathbf{z} e \mathbf{z}' usados na ilustração anterior o incremento Δ é igual a -1 .

Do exposto, o valor da quadrática para $t(\mathbf{z}')$ pode ser escrito como:

$$t(\mathbf{z}') = t(\mathbf{z}) + \Delta t(\mathbf{z}), \quad (4.62)$$

onde:

$t(\mathbf{z})$: é calculado pela (4.59); e

$\Delta t(\mathbf{z})$: devido ao incremento Δ , é calculado pela expressão (JONGE e TIBERIUS, 1996, p. 24):

$$\Delta t(\mathbf{z}) = d_1 \{ 2\Delta [(z_1 - \hat{z}_1) + \sum_{j=2}^t l_{j1} (z_j - \hat{z}_j)] + (\Delta)^2 \}. \quad (4.63)$$

A busca usando a decomposição \mathbf{LDL}^T conforme o procedimento descrito, quando comparada com a forma de organização do método FARA, otimiza a busca nas linhas da tabela, coluna por coluna, porque em cada linha, a variação entre as colunas é na primeira ambigüidade (confirme na tabela da fase (ii) do método FARA).

Do ponto de vista estatístico, é importante guardar além do minimizador ($\tilde{\mathbf{z}}$), solução do problema de mínimos quadrados inteiro transformado (4.55), também, o segundo melhor resultado encontrado ($\tilde{\mathbf{z}}'$).

Por causa das equações em (4.50) e das propriedades da matriz de transformação \mathbf{Z} , existe uma correspondência biunívoca entre as ambigüidades transformadas (\mathbf{z}) e as originais (\mathbf{a}) e, com isso, obtém-se:

$$\tilde{\mathbf{a}} = \mathbf{Z}^{-1} \tilde{\mathbf{z}} \quad \therefore \quad \Omega_2(\tilde{\mathbf{a}}) = (\tilde{\mathbf{a}} - \hat{\mathbf{x}}_2)^T \mathbf{Q}_{\hat{\mathbf{x}}_2}^{-1} (\tilde{\mathbf{a}} - \hat{\mathbf{x}}_2); \quad (4.64)$$

$$\tilde{\mathbf{a}}' = \mathbf{Z}^{-1} \tilde{\mathbf{z}}' \quad \therefore \quad \Omega_2'(\tilde{\mathbf{a}}') = (\tilde{\mathbf{a}}' - \hat{\mathbf{x}}_2)^T \mathbf{Q}_{\hat{\mathbf{x}}_2}^{-1} (\tilde{\mathbf{a}}' - \hat{\mathbf{x}}_2). \quad (4.65)$$

Os algoritmos do método LAMBDA conforme descritos nessa seção, estão todos registrados em JONGE e TIBERIUS (1996). Tais algoritmos, também foram implementados computacionalmente por Kai Borre, com auxílio de Christian Tiberius e encontram-se disponibilizados na lista de M-Files do livro: *Linear Álgebra, Geodesy, and GPS* por STRANG e BORRE (1997).

4.3.5 Solução com Satélites Primários e Secundários

Nesse método, são formados dois grupos de observações. Um primeiro grupo envolvendo 4 satélites, denominados satélites primários, são usados para gerar posições potenciais (baseadas em 3 ambigüidades), e um segundo grupo consistindo dos satélites restantes, denominados satélites secundários, são usados para verificação das posições potenciais. Se a posição baseada nas ambigüidades do primeiro grupo não concorda com as observações/posição envolvendo, também, o segundo grupo, as ambigüidades encontradas são descartadas e um novo conjunto de 3 ambigüidades

deve ser testado.

A formulação matemática do método, com a idéia de usar ajustamento por mínimos quadrados e separar os satélites em grupos primários e secundários, foi descrito por HATCH (1990). Tal método compõe o método integrado descrito por ABIDIN (1992, p.38), ABIDIN (1993).

Uma outra formulação matemática do método, com algumas modificações em relação ao trabalho original de HATCH (1990), é apresentado em LEICK (1995, p. 371). O desenvolvimento a seguir, baseia-se na formulação do último autor.

Suponha-se ter sido rastreado simultaneamente até uma época particular (t)⁸ os mesmos n satélites. As (n-1) equações da dupla diferença de fase podem ser representadas por (ver equação 2.93, negligenciando alguns erros):

$$\left. \begin{aligned}
 \Phi_{AB}^{12}(t) &= \rho_{AB}^{12}(t) - \lambda N_{AB}^{12} \\
 \Phi_{AB}^{13}(t) &= \rho_{AB}^{13}(t) - \lambda N_{AB}^{13} \\
 \Phi_{AB}^{14}(t) &= \rho_{AB}^{14}(t) - \lambda N_{AB}^{14} \\
 \Phi_{AB}^{15}(t) &= \rho_{AB}^{15}(t) - \lambda N_{AB}^{15} \\
 &\vdots \\
 &\vdots \\
 &\vdots \\
 \Phi_{AB}^{ln}(t) &= \rho_{AB}^{ln}(t) - \lambda N_{AB}^{ln}
 \end{aligned} \right\} \begin{array}{l} \text{Grupo de observações DDF dos satélites} \\ \text{primários: 1,2,3,4} \\ \\ \text{Grupo de observações DDF dos satélites} \\ \text{secundários: 5,6,...,n} \end{array} \quad (4.66)$$

A única recomendação feita por HATCH (1990, p. 303) em relação a escolha dos satélites do grupo primário, é que o GDOP seja razoavelmente bom, mas não comenta quanto tem que ser tal valor. Na literatura é comum encontrar recomendações para que o GDOP esteja abaixo de 6 para o posicionamento com as pseudodistâncias. Portanto, abaixo desse valor, o GDOP será considerado razoável.

Obtida a solução flutuante (\hat{x}) do ajustamento de todas as épocas (1,2,...,t) e a

⁸ Nesse método, as ambigüidades são consideradas como quantidades conhecidas e, por isso, não existe deficiência no posto da matriz **A** na primeira época. No entanto, para formar o espaço de busca das ambigüidades, é melhor aproveitar a solução flutuante e a matriz variância-covariância correspondente das ambigüidades provenientes de um ajustamento sequencial envolvendo todas as épocas anteriores à época (t).

correspondente matriz cofatora ($Q_{\hat{i}}$), extrai-se do vetor \hat{x} e da matriz $Q_{\hat{i}}$, as ambigüidades do grupo primário (\hat{x}_p) e a correspondente matriz cofatora ($Q_{\hat{i}_p}$), conforme ilustra a expressão a seguir:

$$\begin{bmatrix}
 x_B \\
 y_B \\
 z_B \\
 N_{AB}^{12} \\
 N_{AB}^{13} \\
 N_{AB}^{14} \\
 N_{AB}^{15} \\
 \vdots \\
 N_{AB}^{1n}
 \end{bmatrix}
 \begin{bmatrix}
 & & & & \\
 & & & & \\
 & & & & \\
 \ddots & & & & \\
 & & & & \\
 & & & & \\
 & & & & \\
 & & & & \\
 & & & & \\
 & & & &
 \end{bmatrix}
 \begin{bmatrix}
 Q_{\hat{i}_B} & Q_{\hat{i}_B \hat{i}_p} & Q_{\hat{i}_B \hat{i}_s} \\
 Q_{\hat{i}_p \hat{i}_B} & Q_{\hat{i}_p} & Q_{\hat{i}_p \hat{i}_s} \\
 Q_{\hat{i}_s \hat{i}_B} & Q_{\hat{i}_s \hat{i}_p} & Q_{\hat{i}_s}
 \end{bmatrix}
 \quad (4.67)$$

O elipsóide de busca das ambigüidades primárias é definido por (ABIDIN, 1993, p. 386):

$$(a_p - \hat{x}_p)^T Q_{\hat{i}_p}^{-1} (a_p - \hat{x}_p) \leq \chi_{3, \alpha}^2, \quad (4.68)$$

onde:

$\chi_{3, \alpha}^2$: percentil da distribuição qui-quadrado com 3 graus de liberdade e nível de significância α ;

\hat{x}_p : vetor com duplas diferenças de ambigüidades estimadas na solução flutuante;

a_p : vetor tri-dimensional representando as ambigüidades inteiras do grupo primário.

A forma linearizada para o sistema (4.66) pode ser escrita como:

$$\mathbf{V}_P = \mathbf{A}_P \mathbf{x} - \mathbf{L}_P ; \quad (4.69)$$

$$\mathbf{V}_S = \mathbf{A}_S \mathbf{x} - \mathbf{L}_S , \quad (4.70)$$

onde:

$\mathbf{x} = [\Delta x_B \ \Delta y_B \ \Delta z_B]^T$: vetor de correções às coordenadas aproximadas da estação incógnita B;

$\mathbf{A}_P \in \mathbf{R}^{3 \times 3}$: matriz das derivadas parciais em relação as coordenadas aproximadas da estação incógnita para o grupo primário;

$\mathbf{A}_S \in \mathbf{R}^{(n-4) \times 3}$: matriz das derivadas parciais em relação as coordenadas aproximadas da estação incógnita para o grupo secundário;

$\mathbf{L}_P = [\Phi_{AB(P)} - \rho_{AB(P)} + \lambda \mathbf{a}_P]$: vetor do grupo primário;

$\mathbf{L}_S = [\Phi_{AB(S)} - \rho_{AB(S)} + \lambda \mathbf{a}_S]$: vetor do grupo secundário.

Com as ambigüidades (\mathbf{a}_P) definidas pelo elipsóide (4.68), calcula-se o vetor \mathbf{L}_P e obtém-se uma solução para \mathbf{x} na equação (4.69), por:

$$\mathbf{x}_P = (\mathbf{A}_P^T \mathbf{Q}_{L_P}^{-1} \mathbf{A}_P)^{-1} \mathbf{A}_P^T \mathbf{Q}_{L_P}^{-1} \mathbf{L}_P. \quad (4.71)$$

Com o vetor \mathbf{x}_P , calcula-se as coordenadas da estação incógnita B, isto é, $\mathbf{x}_B = [x_B \ y_B \ z_B]^T$. Agora, procedendo como no método da linha de base conhecida (ver seção 4.12, equação 4.4), calcula-se as ambigüidades do grupo secundário:

$$N_{AB(S)}^{1,i} = \frac{1}{\lambda} (\rho_{AB}^{1,i}(t) - \Phi_{AB}^{1,i}(t)), \ i = 5, 6, \dots, n. \quad (4.72)$$

Os valores encontrados para $N_{AB(S)}^{1,i}$ devem ser arredondados para seu inteiro mais próximo, obtendo-se desta forma, o vetor de ambigüidades do grupo secundário

$$\mathbf{a}_S = \text{round} (N_{AB(S)}). \quad (4.73)$$

Usando, \mathbf{a}_S , para calcular o vetor \mathbf{L}_S em (4.70), estima-se um novo (\mathbf{x}_c)

incluindo as observações do primeiro e segundo grupo, como:

$$\mathbf{x}_C = (\mathbf{A}_C^T \mathbf{Q}_L^{-1} \mathbf{A}_C)^{-1} \mathbf{A}_C^T \mathbf{Q}_L^{-1} \mathbf{L}, \quad (4.74)$$

onde:

$$\mathbf{A}_C = \begin{bmatrix} \mathbf{A}_P \\ \mathbf{A}_S \end{bmatrix};$$

$$\mathbf{L} = \begin{bmatrix} \mathbf{L}_P \\ \mathbf{L}_C \end{bmatrix}.$$

Estando as ambigüidades (\mathbf{a}_P) do grupo primário corretas, as ambigüidades do grupo secundário (\mathbf{a}_S) também estarão e, por causa da precisão das medidas da fase da portadora, a diferença entre os vetores calculados deve ser pequena:

$$\Delta \mathbf{x} = \|\mathbf{x}_C - \mathbf{x}_P\|_2 \leq \varepsilon, \quad (4.75)$$

com $\varepsilon < 10$ cm, por exemplo.

Os vetores ($\mathbf{a} = [\mathbf{a}_P \ \mathbf{a}_S]^T$) que fizerem com que se cumpra a condição (4.75) (e particularmente aquele que fornecer $\|\Delta \mathbf{x}\| = \text{mínimo}$) serão os candidatos a solução ($\check{\mathbf{a}} = [\check{\mathbf{a}}_P \ \check{\mathbf{a}}_S]^T$), devendo portanto, serem cuidadosamente analisados. Inclui-se nesta análise, os testes de hipótese: $H_0 : \check{\sigma}_0^2 = \sigma_0^2$; e $H_0 : \check{\sigma}_0^2 = \sigma_0^2$, (ver seção 4.3.2).

O método exposto pode ser aplicado época, por época, ou pode ser aplicado em um ajustamento seqüencial.

4.4 MÉTODOS INTEGRADOS

Os métodos integrados de resolução de ambigüidades, de acordo com HAN e RIZOS (1998, p. 39), combinam os métodos de procura no espaço das ambigüidades, no espaço das coordenadas e, quando os receptores rastreiam nas duas frequências, utilizam as observáveis formadas das combinações lineares das portadoras L_1 e L_2 .

Aplicam, ainda, ajustamento das pseudodistâncias suavizadas pela fase da portadora e quando possível, combinam dados de outros sensores, tais como: sistema de navegação inercial, bússolas, etc.

ABIDIN (1992), foi um dos primeiros a utilizar métodos integrados para resolver as ambigüidades OTF. Em seu trabalho, foram utilizados: pseudodistâncias suavizadas pela fase da portadora, combinação linear extra-widelaning, busca no espaço das coordenadas usando a função de ambigüidade, busca no espaço das ambigüidades usando os grupos de satélites primários e secundários (este método é também conhecido por busca no espaço das ambigüidades por mínimos quadrados) e, além disso, várias estações de coordenadas conhecidas denominadas de estações de monitoramento.

Nesse trabalho, no entanto, qualquer método que integre dois ou mais métodos envolvendo conceitos de busca em espaços diferentes (exemplo: espaço das coordenadas e espaço das ambigüidades) será considerado um método integrado. Nesse sentido, dois métodos integrados serão apresentados: o primeiro, fundamenta-se na teoria dos métodos de funções de penalidade para minimização de funções não-lineares com injunções não-lineares de desigualdade, procurando resolver o problema de mínimos quadrados misto, isto é, quando parte dos parâmetros são reais e parte são inteiros; o segundo, utiliza buscas direcionais relacionadas no espaço das ambigüidades e no espaço das coordenadas, explorando também, os grupos de satélites primários e secundários.

4.4.1 O Problema de Mínimos Quadrados Linear Misto

No método FARA (seção 4.3.2) o fator de variância a posteriori ($\hat{\sigma}_0^2$) (equação 4.34) é calculado usando o vetor dos resíduos (\check{V}) cuja forma quadrática, $\Omega = \check{V}^T Q_L^{-1} \check{V}$, (equação 4.35) é mínima. \check{V} é obtido injuncionando no ajustamento o vetor de ambigüidades $\check{x}_2 = \check{a}$. Foi demonstrado que Ω pode ser calculado

separadamente pela soma $\Omega_1 + \Omega_2$ (equação 4.40), com Ω_1 o valor da forma quadrática da solução flutuante do problema de mínimos quadrados e Ω_2 vem da solução do problema de mínimos quadrados inteiro (equação 4.64).

No problema de mínimos quadrados misto (ou problema de mínimos quadrados linear misto, quando o valor aproximado (\mathbf{x}_0) para a linearização estiver bem próximo da solução), espera-se encontrar $\tilde{\mathbf{x}} = [\tilde{\mathbf{x}}_1 \quad \tilde{\mathbf{a}}]^T$ de modo que a quadrática $\Omega = \Omega_1 + \Omega_2$ seja mínima e, assim, idêntica as soluções encontradas pelos outros métodos.

O problema de mínimos quadrados linear misto pode ser enunciado como:

$$\underset{\mathbf{x} = \begin{bmatrix} \mathbf{x}_1 \in \mathbb{R}^{n_1} \\ \mathbf{x}_2 \in \mathbb{Z}^{n_2} \end{bmatrix}}{\text{minimizar}} \quad \Omega(\mathbf{x}) = \|\mathbf{Ax} - \mathbf{L}\|_{Q_L^{-1}}^{-1}. \quad (4.76)$$

Os métodos propostos para resolver o problema (4.76), baseados nos conceitos de minimização sem injunções e funções de penalidade, são desenvolvidos nas próximas subseções.

4.4.1.1 Métodos de funções de penalidade

Os métodos de função de penalidade são métodos clássicos da Programação Não-Linear para resolver o seguinte tipo de problema:

$$\begin{aligned} &\underset{\mathbf{x} \in \mathbb{R}^n}{\text{minimizar}} \quad f(\mathbf{x}) \\ &\text{sujeito a: } g_i(\mathbf{x}) \geq 0, \quad i = 1, \dots, t \end{aligned} \quad (4.77)$$

onde:

$f: \mathbb{R}^n \rightarrow \mathbb{R}$: não-linear;

$g_i: \mathbb{R}^n \rightarrow \mathbb{R}$: não-linear.

O problema (4.77) é, então, trocado por um problema transformado sem injunções:

$$\underset{\mathbf{x} \in \mathbf{R}^n}{\text{minimizar}} \quad \phi(\mathbf{x}, r_k) = f(\mathbf{x}) + r_k \sum_{i=1}^r G_i[g_i(\mathbf{x})], \quad (4.78)$$

onde:

$r_k > 0$: parâmetro de penalidade;

$r_k \sum_{i=1}^r G_i[g_i(\mathbf{x})]$: termo de penalidade;

$G_i[g_i(\mathbf{x})]$: função de penalidade.

Se a busca ao vetor solução \mathbf{x}^* minimizador de (4.78), é escolhida para ser só dentro da região factível, deve-se iniciar em um ponto \mathbf{x}_0 de tal modo que nenhuma injunção $g_i(\mathbf{x}_0)$ seja violada. A função G_i pode ser definida para ser:

$$G_i = \frac{1}{g_i(\mathbf{x})}; \quad \text{ou} \quad (4.79)$$

$$G_i = \log[g_i(\mathbf{x})]. \quad (4.80)$$

Usando-se as funções (4.79) ou (4.80), toda vez que durante a busca alguma das injunções for ameaçada de violação, ou seja, $g_i \rightarrow 0$ implicaria em aumento de $G_i \rightarrow \infty$.

Os métodos que trabalham dessa forma usando funções como as (4.79) ou (4.80) são chamados métodos de penalidade interior ou métodos de barreira.

Por outro lado, muitas vezes não se conhece um ponto \mathbf{x}_0 interior (ponto da região factível) para iniciar o processo iterativo e usar o método de barreira. Assim sendo, um outro tipo de função pode ser usada, por exemplo:

$$G_i = \{\max[0, -g_i(\mathbf{x})]\}^q, \quad \text{com } q > 1. \quad (4.81)$$

Quando são usadas funções como as (4.81), existe a penalização para a função $\phi(\mathbf{x}, r_k)$, toda vez que uma ou mais das injunções forem violadas. Os métodos que utilizam funções como a (4.81) são chamados métodos de penalidade exterior ou, simplesmente, métodos de penalidades.

Com relação ao parâmetro de penalidade r_k , no método de barreira, ele deve começar com um valor razoavelmente alto (exemplo: $r_1=1.000$ ou $r_1=10.000$) e gradativamente ser decrescido ($r_{k+1} = 0,1 r_k$). Se para a sequência de valores de r_k , $k=1,2, \dots$, a função $\phi(\mathbf{x}, r_k)$ é minimizada, então, quanto r_k tende para 0, o mínimo da função $\phi(\mathbf{x}, r_k)$ tende para o mínimo da função $f(\mathbf{x})$ com as injunções $g_i(\mathbf{x})$, $i=1,\dots,t$.

No método da função de penalidade, o parâmetro r_k deve começar com um valor razoavelmente pequeno (exemplo: $r_k = 0,1$ ou $r_k = 0,001$) e gradativamente ser aumentado ($r_{k+1} = 10 r_k$). Se a função $\phi(\mathbf{x}, r_k)$ é minimizada para a sequência de valores r_k , $k=1,2, \dots$, então, quando r_k tende para o infinito, o mínimo de $\phi(\mathbf{x}, r_k)$ tende para o mínimo de $f(\mathbf{x})$ com as injunções $g_i(\mathbf{x})$, $i=1,\dots,t$.

A prova das convergências mencionadas nos parágrafos anteriores para os métodos de barreira e penalidade podem ser encontradas, por exemplo, em: RAO (1979, p.392,403), BAZARAA e SHETTY (1979, p. 346, 338).

Pertencente a família dos métodos de penalidade, investigou-se também nesse trabalho, o método da penalidade hiperbólica devido a XAVIER (1982)⁹, apud XAVIER (1998).

A função de penalidade hiperbólica é definida como:

$$G(y, \alpha, \tau) = -(\frac{1}{2} \operatorname{tg} \alpha) y + \sqrt{(\frac{1}{2} \operatorname{tg} \alpha)^2 y^2 + \tau^2}, \quad (4.82)$$

onde:

$$\alpha \in [0, \pi/2] ; e$$

$$\tau \geq 0 .$$

Uma forma alternativa, porém, mais conveniente de usá-la é:

⁹ XAVIER, A. E. Penalização hiperbólica: um novo método para resolução de problemas de otimização. Dissertação (Mestrado). Universidade Federal do Rio de Janeiro. COPPE, Rio de Janeiro, 1982.

$$G(y, \lambda, \tau) = -\lambda y + \sqrt{\lambda^2 y^2 + \tau^2}, \quad (4.83)$$

onde:

$$\lambda \geq 0; \text{ e}$$

$$\tau \geq 0.$$

A função de penalidade hiperbólica tem características muito especiais e úteis para aplicação dos métodos de minimização, são elas (XAVIER, 1998, p. 3,4):

1. $G(y, \lambda, \tau)$ é k-vezes diferenciável continuamente para qualquer inteiro positivo k para $\tau > 0$;
2. $G(y, \lambda, \tau)$ é assintoticamente tangente a linha $r_1(y) = -2\lambda y$ e $r_2(y) = 0$ para $\tau > 0$;
3. $G(y, \lambda, \tau) \geq -2\lambda y$ para todo $y, \lambda \geq 0, \tau \geq 0$;
4. $G(0, \lambda, \tau) = \tau$ para $\tau \geq 0$ e $\lambda \geq 0$;
5. $G(y, \lambda, \tau)$:
 - i) é uma função convexa decrescente em y para $\tau > 0$ e $\lambda \geq 0$;
 - ii) é uma função convexa não-crescente em y para $\tau = 0$ e $\lambda \geq 0$;
 - iii) é uma função convexa igual a τ para $\lambda = 0$;
6. Para $\lambda^{k+1} > \lambda^k$ e $\tau > 0$:
 - i) $G(y, \lambda^{k+1}, \tau) < G(y, \lambda^k, \tau)$ para $y > 0$;
 - ii) $G(y, \lambda^{k+1}, \tau) < G(y, \lambda^k, \tau) = \tau$ para $y = 0$;
 - iii) $G(y, \lambda^{k+1}, \tau) < G(y, \lambda^k, \tau)$ para $y < 0$;
7. $G(y, \lambda, \tau^{k+1}) < G(y, \lambda, \tau^k)$ para todo $y, \lambda > 0, 0 \leq \tau^{k+1} < \tau^k$;
8. $\max_y [G(y, \lambda, \tau^0) - G(y, \lambda, \tau^1)] = \tau^0 - \tau^1$ e ocorre em $y = 0$ para $\lambda > 0$ e $0 \leq \tau^1 < \tau^2$.

A derivada da função de penalidade hiperbólica em relação a y assume a forma:

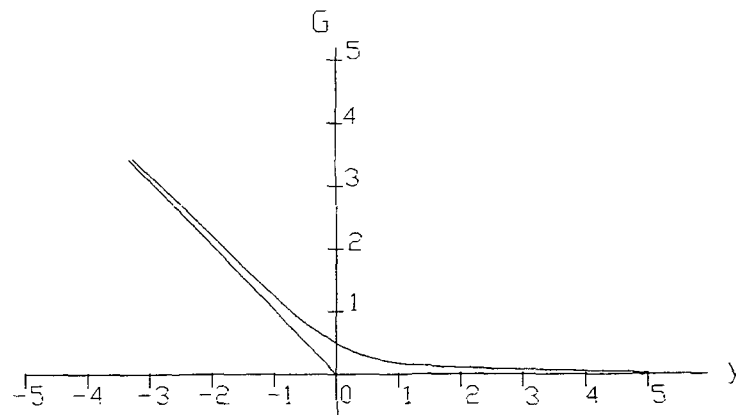
$$G'(y) = \lambda [-1 + \lambda y / \sqrt{\lambda^2 y^2 + \tau^2}], \quad (4.84)$$

e tem as seguintes propriedades:

9. $G'_y(y, \lambda, \tau) = -\lambda$ para $\tau > 0$;
10. $G'_y(0, \lambda, \tau) = -\lambda$ para $\tau > 0$;
11. quando o parâmetro λ aumenta a derivada da função de penalidade $G'_y(y, \lambda, \tau)$ decresce para os pontos $y < \bar{y}$ e aumenta para os pontos $y > \bar{y}$ onde $\bar{y} = \beta\tau/\lambda$ com $\beta = \sqrt{(-1 + \sqrt{5})/2}$

O gráfico da função de penalidade é ilustrado na figura (4.3).

FIGURA 4.3 – GRÁFICO DA FUNÇÃO DE PENALIDADE HIPERBÓLICA



FONTE: XAVIER A. E. **Hyperbolic penalty: new method for nonlinear programming with inequalities.** 1998

XAVIER (1998), apresenta o algoritmo 4.1, para aplicação da função de penalidade hiperbólica na minimização do problema transformado:

$$\underset{\mathbf{x} \in \mathbb{R}^n}{\text{minimizar}} \quad \phi(\mathbf{x}, \lambda^k, \tau^k) = f(\mathbf{x}) + \sum_{i=1}^l G_i[g_i(\mathbf{x}), \lambda^k, \tau^k]. \quad (4.85)$$

Algoritmo 4.1 (aplicação da função de penalidade hiperbólica)

1. faça $k = 0$, tome valores iniciais \mathbf{x}^0 , $\lambda^1 > 0$ e $\tau^1 > 0$;
2. faça $k = k + 1$ e resolva o problema de minimização sem injunção

$$\min \phi(\mathbf{x}, \lambda^k, \tau^k)$$
 a partir do ponto \mathbf{x}^{k-1} e obtenha o ponto ótimo \mathbf{x}^k ;

3. teste de factibilidade
se \mathbf{x}^k é um ponto infactível então execute o passo 4 senão execute o passo 5;
4. decresça o parâmetro λ :
 $\lambda_i^{k+1} = r \lambda_i^k, r > 1$, vá para o passo 2;
5. decresça o parâmetro τ :
 $\tau_i^{k+1} = q \tau_i^k, 0 < q < 1$, vá para o passo 2.

Detalhes a respeito da convergência do método são encontrados nas referências de XAVIER (1998) e (1982), citadas no texto.

4.4.1.2 Métodos de funções de penalidade sobre os não-inteiros

Quando se deseja estimar alguns ou todos os parâmetros sobre os inteiros, pode-se usar uma função para penalizar uma busca toda vez que o valor para o parâmetro desejado não for inteiro.

Particionando o vetor $\mathbf{x} = [\mathbf{x}_1 \ \mathbf{x}_2]^T$, onde $\mathbf{x}_1 \in \mathbf{R}^{n_1}$ e $\mathbf{x}_2 \in \mathbf{Z}^{n_2}$, o problema transformado (4.76) pode ser colocado na forma (RAO, 1979, p. 555):

$$\text{minimizar } \phi(\mathbf{x}, r_k, s_k) = f(\mathbf{x}) + r_k \sum_{i=1}^t G_i[g_i(\mathbf{x})] + s_k Q_k(\mathbf{x}_2) \quad (4.86)$$

onde:

$r_k > 0$: parâmetro de penalidade;

$r_k \sum_{i=1}^t G_i[g_i(\mathbf{x})]$: termo de penalidade para as injunções;

s_k : parâmetro de penalidade para os parâmetros \mathbf{x}_2 não-inteiros;

$s_k Q_k(\mathbf{x}_2)$: termo de penalidade a ser ativado sempre que os valores em \mathbf{x}_2 não forem inteiros;

$Q_k(\mathbf{x}_2)$: função de penalidade para os valores em \mathbf{x}_2 não-inteiros;

A função $Q_k(\mathbf{x}_2)$ deve possuir as seguintes características:

$$Q_k(\mathbf{x}_2) = \begin{cases} 0 & \text{se } \mathbf{x}_2 \in \mathbf{Z}^{n_2} \\ u > 0 & \text{se } \mathbf{x}_2 \notin \mathbf{Z}^{n_2}, \text{ com } u \in \mathbf{R} \end{cases} \quad (4.87)$$

e pode ser definida, por exemplo, como:

$$Q_k(\mathbf{x}_2) = \sum_{x_i \in \mathbf{x}_2} \left\{ 4 \left(\frac{x_i - y_i}{z_i - y_i} \right) \left(1 - \frac{x_i - y_i}{z_i - y_i} \right) \right\}^{\beta_k}, \quad (4.88)$$

onde:

$y_i \leq x_i$, com $y_i = \text{int}_{\text{esquerda}}(x_i)$: inteiro a esquerda de x_i ;

$z_i \geq x_i$, com $z_i = \text{int}_{\text{direita}}(x_i)$: inteiro a direita de x_i ;

$\beta_k \geq 1$.

A escolha dos valores iniciais para os parâmetros de penalidade r_k , s_k e também para β_k , influenciam na razão de convergência do método e, grande parte das vezes, só com experiências práticas é que se pode definir um valor ótimo para eles.

Quando não se conhecem esses valores a priori, pode-se determiná-los segundo as expressões (RAO, 1979, p. 557-559):

$$s_1 = c_1 \frac{P'_1(\mathbf{x}_{(1)}, r_1)}{Q'_1(\mathbf{x}_{(1)}, \beta_1)}, \quad (4.89)$$

onde:

$$P'_k = \frac{1}{\sqrt{n}} \|\nabla P_k^T\|_2, \quad (4.90)$$

com:

$$P_k = f(\mathbf{x}) + r_k \sum_{i=1}^l G_i[g_i(\mathbf{x})]; \quad (4.91)$$

$$Q'_k = \frac{1}{2} 4^{\beta_k} \beta_k (\beta_k - 1) \beta^{k-1} (2\beta_{k-1})^{\frac{1}{2} - \beta_k}; \quad (4.92)$$

c_1 : é escolhido com os valores 0,001 ou 0,1;

$$r_1 = c_2 \frac{f(\mathbf{x}_1)}{\sum_{i=1}^l G_i[g_i(\mathbf{x})]}, \quad c_2 = 10 c_1; \quad (4.93)$$

β_1 deve ser escolhido a fim de manter a continuidade das derivadas primeiras de (4.88), um valor poderia ser, por exemplo: $\beta_1 = 2,2$;

$\mathbf{x}_{(1)}$ é a aproximação inicial do vetor de parâmetros \mathbf{x} .

A atualização do parâmetro r_k é feita por:

$$r_{k+1} = c_3 r_k, \quad (4.94)$$

onde c_3 depende do tipo de função usada (barreira ou penalidade).

A atualização do parâmetro s_k é feita por:

$$s_{k+1} = c_4 s_k, \quad \text{com } 1,4 < c_4 < 4,5 \quad (4.95)$$

e a atualização do parâmetro β_k é:

$$\beta_{k+1} = c_5 \beta_k, \quad \text{com } 0,7 < c_5 < 0,9. \quad (4.96)$$

Maiores detalhes com discussões de ordem prática desse método são encontradas em RAO (1979, p. 554-562).

4.4.1.3 Solução do problema de mínimos quadrados linear misto por métodos de penalidade

A idéia inicial de resolver as ambigüidades por meio da solução do problema de mínimos quadrados misto (4.76), aplicando métodos de penalidade, foi apresentada primeiramente em GARNÉS (1998_(c)). Nessa seção uma estratégia para tal proposta é desenvolvida.

Formulação do Problema

Suponha-se ter sido realizado um ajustamento com as duplas diferenças de fase até a época $t > 1$, e que o vetor dos parâmetros da solução flutuante ($\hat{\mathbf{x}}$) foi determinado, assim como, a correspondente matriz variância-covariância ($\Sigma_{\hat{\mathbf{x}}} = \hat{\sigma}_0^2 Q_{\hat{\mathbf{x}}}$), como mostra a equação (4.66). Das variâncias, elementos da diagonal de $\Sigma_{\hat{\mathbf{x}}}$, determina-se o desvio padrão de cada uma das ambigüidades estimadas:

$$\begin{aligned} \hat{N}_{AB}^{12} &, & \sigma_{\hat{N}_{AB}^{12}} \\ \hat{N}_{AB}^{13} &, & \sigma_{\hat{N}_{AB}^{13}} \end{aligned}$$

η : é um número inteiro para definir o tamanho do intervalo (um bom valor é $\eta = 3$ ou $\eta = 4$);
 $\sigma_{\hat{x}_i}$, $i = 4, \dots, n_2$: é o desvio padrão das ambigüidades estimadas pela solução flutuante.

Conforme descrito na seção (4.4.1.2), o problema (4.99) com injunções, pode ser transformado no problema sem injunções usando as funções de penalidade G_i e Q_k , como:

$$\text{minimizar } \phi(\mathbf{x}, \mathbf{r}_k, \mathbf{s}_k) = \Omega(\mathbf{x}) + \mathbf{r}_k^T \sum_{i=1}^{i=2n_2} G_i[g_i(\mathbf{x})] + \mathbf{s}_k^T Q_k(\mathbf{x}_2), \quad (4.100)$$

ou, com a função de penalidade hiperbólica, como:

$$\text{minimizar } \phi(\mathbf{x}, \lambda^k, \tau^k) = \Omega(\mathbf{x}) + \sum_{i=1}^{i=2n_2} G_i[g_i(\mathbf{x}), \lambda^k, \tau^k] + \mathbf{s}_k^T Q_k(\mathbf{x}_2). \quad (4.101)$$

Na minimização dos problemas (4.100) e (4.101) pode-se aplicar, por exemplo, o algoritmo 3.11 (Newton modificado globalizado com a decomposição de Cholesky modificada).

Os parâmetros \mathbf{r}_k , \mathbf{s}_k , β_k , λ_k e τ_k são atualizados conforme visto nas seções anteriores.

O método proposto, quando converge, encontra um ponto que satisfaz o requerimento: $\mathbf{x}_2 \in \mathbf{Z}^{n_2}$.

4.4.2 Busca Direcional nos Espaços das Ambigüidades e das Coordenadas

A idéia de utilizar buscas direcionais para resolver as ambigüidades surgiu quando se testava a minimização inteira na direção de máxima descida ($-\nabla f(\mathbf{a})$) da quadrática bi-dimensional $f(\mathbf{a}) = (\mathbf{a} - \hat{\mathbf{x}})^T \mathbf{Q}_i^{-1} (\mathbf{a} - \hat{\mathbf{x}})$, onde aqui, $\hat{\mathbf{x}}$ representa o vetor bi-dimensional das ambigüidades reais estimadas; e \mathbf{a} o vetor incógnito bi-dimensional de inteiros.

Para o esclarecimento da idéia, será usado um exemplo prático retirado da 12ª época do teste para linha de base 2 (P003 – PC) (ver anexo 2).

Nesta época, duas das componentes das ambigüidades reais estimadas ($\hat{\mathbf{x}}$), com a correspondente submatriz cofatora ($\mathbf{Q}_{\hat{\mathbf{x}}}$), foram:

$$\hat{\mathbf{x}} = \begin{bmatrix} -172\,710,649\,44 \\ 115\,806,049\,47 \end{bmatrix}; \quad \mathbf{Q}_{\hat{\mathbf{x}}} = \begin{bmatrix} 1,6101114069 & -0,9569706649 \\ -0,9569706649 & 0,6110010269 \end{bmatrix}.$$

A inversa de $\mathbf{Q}_{\hat{\mathbf{x}}}$ é:

$$\mathbf{Q}_{\hat{\mathbf{x}}}^{-1} = \begin{bmatrix} 8,987045 & 14,075816 \\ 14,075816 & 23,682682 \end{bmatrix}.$$

A direção de máxima descida da quadrática $f(\mathbf{a})$, $\mathbf{d}_k = -\nabla f(\mathbf{a})$, é :

$$\mathbf{d}_k = -\nabla f(\mathbf{a}) = 2 \mathbf{Q}_{\hat{\mathbf{x}}}^{-1} (\hat{\mathbf{x}} - \mathbf{a}), \quad (4.102)$$

que pode ser normalizada por:

$$\bar{\mathbf{d}}_k = \frac{\mathbf{d}_k}{\|\mathbf{d}_k\|}. \quad (4.103)$$

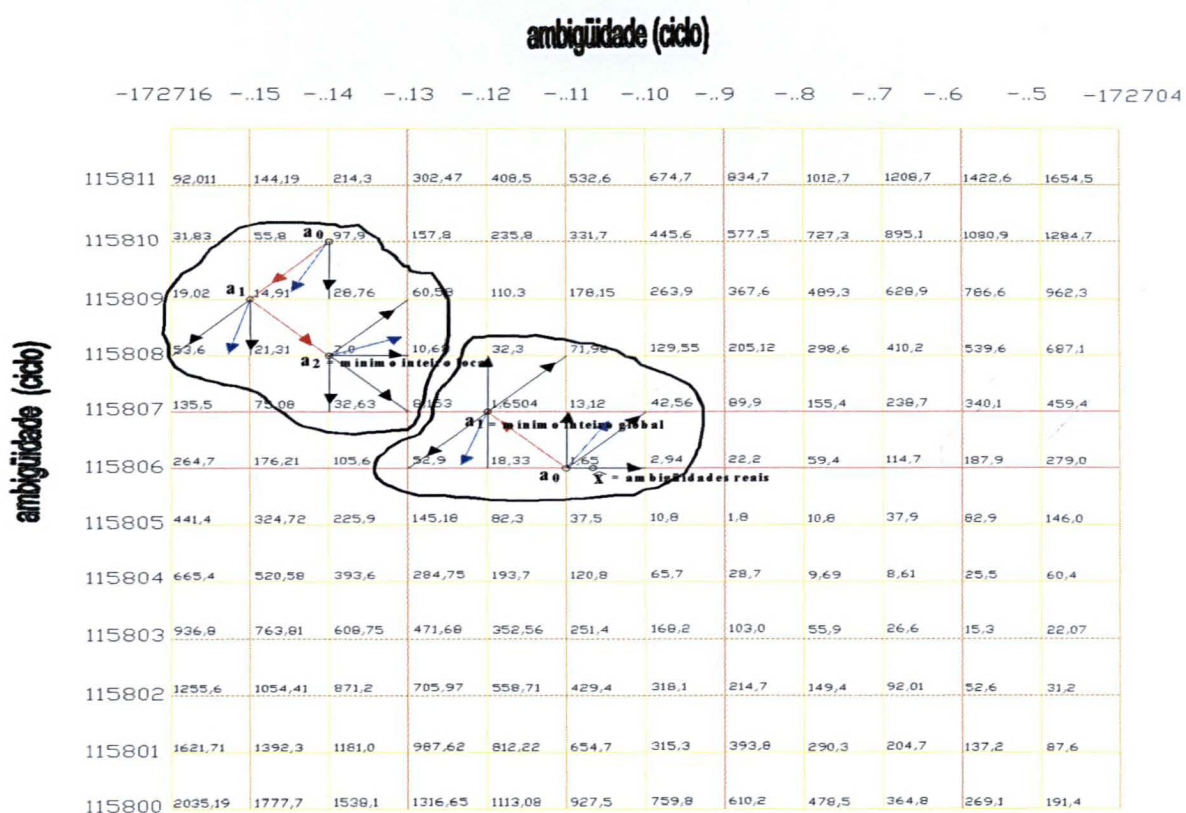
O método que busca a minimização inteira a partir da direção $\bar{\mathbf{d}}_k$, começa em um ponto inicial \mathbf{a}_0 e procura quais pontos inteiros estão mais próximos da direção normalizada $\bar{\mathbf{d}}_{k_0}$ (direção em \mathbf{a}_0). Aquele ponto que fornecer o menor valor da quadrática $f(\mathbf{a})$ é atualizado para ser \mathbf{a}_1 . Nele, calcula-se uma nova direção normalizada $\bar{\mathbf{d}}_{k_1}$ e a partir dela encontra-se o novo ponto inteiro próximo da direção $\bar{\mathbf{d}}_{k_1}$ que minimiza $f(\mathbf{a})$. Faz-se a atualização para este ponto e chama-o de \mathbf{a}_2 . O processo de minimização continua até que um ponto inteiro, mínimo local, seja encontrado. Isto significa, mesmo que exista uma direção de descida a partir de tal ponto, não existe nenhum vizinho, inteiro, que forneça menor valor para forma quadrática $f(\mathbf{a})$.

Quando um mínimo local inteiro é encontrado, deve-se escolher outro ponto inicial \mathbf{a}_0 e reiniciar o processo de minimização.

Se o espaço onde foram encontrados os mínimos locais, for suficientemente grande, de modo que o mínimo global esteja contido nele, então, o método encontra o minimizador ($\tilde{\mathbf{a}}$) da quadrática $f(\tilde{\mathbf{a}}) = (\tilde{\mathbf{a}} - \hat{\mathbf{x}})^T \mathbf{Q}_x^{-1} (\tilde{\mathbf{a}} - \hat{\mathbf{x}})$. Além disso, basta guardar os mínimos locais encontrados para se ter o segundo menor valor ($\tilde{\mathbf{a}}'$) da quadrática.

Na figura 4.4 foi gerado um grille de pontos inteiros com as respectivas cotas de $f(\mathbf{a})$ para o exemplo colocado. As setas em azul indicam as direções de máxima descida da função $f(\mathbf{a})$ no ponto considerado, as setas em preto indicam as direções dos vizinhos inteiros e as setas em vermelho indicam a direção do vizinho inteiro que fornece o menor valor para $f(\mathbf{a})$. A linha em torno das setas limita a região onde o método converge para o mesmo ponto.

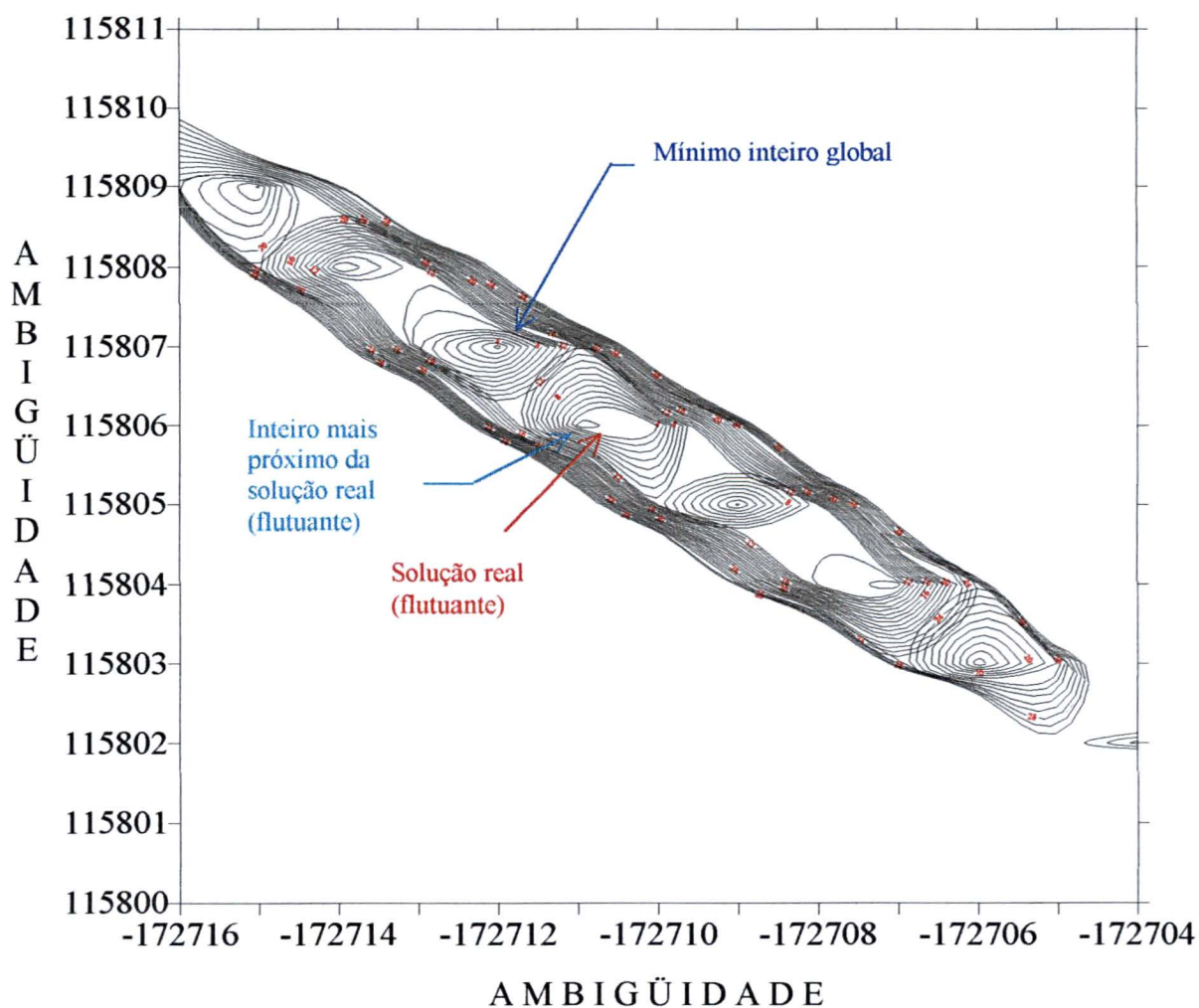
FIGURA 4.4 – BUSCA INTEIRA PELA MÁXIMA DIREÇÃO DE DESCIDA



NOTA : as linhas limitando o contorno de convergência estão representadas apenas aproximadamente.

Analisando o comportamento das cotas de $f(\mathbf{a})$, percebeu-se que os valores mínimos ficavam aproximadamente alinhados segundo uma determinada direção. Plotando-se as curvas de nível da função, comprovou-se o fato, conforme mostra a figura 4.5.

FIGURA 4.5 – ESPAÇO DE BUSCA DAS AMBIGÜIDADES: CASO BI-DIMENSIONAL



Com o vetor das ambigüidades reais estimadas ($\hat{\mathbf{x}} = [-172710, 64944 \ 115806, 04947]^T$) e com o vetor de ambigüidades inteiras ($\tilde{\mathbf{a}}$) que minimiza a forma quadrática $f(\mathbf{a})$, calculou-se a direção a partir de $\hat{\mathbf{x}}$, para a qual o ponto $\tilde{\mathbf{a}}$ se encontra.

A direção foi:

$$\mathbf{d} = \mathbf{\check{a}} - \hat{\mathbf{x}} = \begin{bmatrix} -172712 - (-172710,6494) \\ 115807 - (115806,04947) \end{bmatrix} = \begin{bmatrix} -1,35056 \\ 0,95053 \end{bmatrix}. \quad (4.104)$$

Normalizando a direção \mathbf{d} , tem-se:

$$\bar{\mathbf{d}} = \begin{bmatrix} -0,817767 \\ 0,575548 \end{bmatrix}. \quad (4.105)$$

Sabe-se que os semi-eixos da elipse (no caso bi-dimensional ou hiper-elipsóide para dimensões acima do \mathbf{R}^3):

$$(\mathbf{a} - \hat{\mathbf{x}})^T \mathbf{Q}_i^{-1} (\mathbf{a} - \hat{\mathbf{x}}) = c^2, \quad (4.106)$$

são obtidos por:

$$c\sqrt{\lambda_i} \mathbf{e}_i, \quad (4.107)$$

sendo:

c^2 : constante que define o tamanho da elipse;

λ_i : i-ésimo autovalor de \mathbf{Q}_i ;

\mathbf{e}_i : i-ésimo autovetor de \mathbf{Q}_i associado ao autovalor λ_i , define a direção do i-ésimo semi-eixo da elipse;

$c\sqrt{\lambda_i}$: comprimento do i-ésimo semi-eixo.

Para o exemplo colocado, os comprimentos e as respectivas orientações dos semi-eixos da elipse, com $c^2=1$, são:

$$\sqrt{\lambda_1} = 1,4799 \quad \therefore \quad \mathbf{e}_1 = [-0,8552 \quad 0,5183]^T; \quad (4.108)$$

$$\sqrt{\lambda_2} = 0,1762 \quad \therefore \quad \mathbf{e}_2 = [-0,5183 \quad -0,8552]^T. \quad (4.109)$$

Observa-se que a direção do semi-eixo maior da elipse (\mathbf{e}_1) com a direção $\bar{\mathbf{d}}$ tem um ângulo (θ) de:

$$\theta = \arccos \left(\frac{\langle \mathbf{e}_1, \bar{\mathbf{d}} \rangle}{\|\mathbf{e}_1\| \|\bar{\mathbf{d}}\|} \right) \cong 3^\circ 55'. \quad (4.110)$$

Com base na distribuição qui-quadrado, com $\nu=2$ graus de liberdade e $\alpha=5\%$, define-se o valor $c^2=7,38$ para uma elipse de confiança com probabilidade $P=97,5\%$ (ver GEMAEL, 1994, p. 243). Para essa elipse o comprimento do semi-eixo maior é:

$$\sigma_{MAX} = c \sqrt{\lambda_{MAX}} = 2,72(1,4799) = 4,020 \text{ ciclos.} \quad (4.111)$$

A diferença entre as duas direções \mathbf{e}_1 e $\bar{\mathbf{d}}$, no final desse comprimento será de:

$$\Delta d = \operatorname{tg} \theta \sigma_{MAX} = 0,27 \text{ ciclos} \quad (4.112)$$

e, portanto, pode-se usar $\mathbf{e}_{MAX} \approx \bar{\mathbf{d}}$.

O resultado do desenvolvimento acima mostra que o mínimo inteiro ($\tilde{\mathbf{a}}$), da forma quadrática $f(\mathbf{a})$, pode ser encontrado realizando uma busca direcional, inteira, na direção do semi-eixo maior da elipse, a partir do seu centro ($\hat{\mathbf{x}}$). Os vetores de ambigüidades candidatos (\mathbf{a}), são definidos por:

$$\mathbf{a}_k = \operatorname{round}(\hat{\mathbf{x}} + t_k \mathbf{e}_{MAX}), \quad (4.113)$$

com:

$$t_k = 0, \pm 0,5, \pm 1, \pm 1,5, \dots, \pm \sigma_{MAX};$$

round : arredonda cada componente do vetor para seu valor inteiro mais próximo.

Observa-se que para o exemplo colocado, com $t_k=1$, o valor \mathbf{a}_k em (4.113) será:

$$\begin{aligned}
\mathbf{a}_k &= \text{round} \left(\begin{bmatrix} -172710,64944 \\ 115806,04947 \end{bmatrix} + 1 \begin{bmatrix} -0,8552 \\ 0,5183 \end{bmatrix} \right) \\
&= \text{round} \left(\begin{bmatrix} -172711,5046 \\ 115806,5678 \end{bmatrix} \right) = \begin{bmatrix} -172712 \\ 115807 \end{bmatrix},
\end{aligned} \tag{4.114}$$

que é o mínimo inteiro da quadrática $f(\mathbf{a})$, conforme mostra a fig.(4.4) e fig.(4.5).

A metodologia apresentada até aqui sugere; ao invés de realizar uma busca a todos os candidatos interiores do espaço (4.106), bastaria encontrar qual dos candidatos definidos por (4.113), fornece o menor valor da forma quadrática $f(\mathbf{a})$ para se ter mínimo inteiro global.

A idéia desenvolvida anteriormente para o espaço bi-dimensional é, agora, generalizada para o espaço elipsoidal t-dimensional:

$$(\mathbf{a} - \hat{\mathbf{x}})^T \mathbf{Q}_{\hat{\mathbf{x}}}^{-1} (\mathbf{a} - \hat{\mathbf{x}}) = \chi^2, \tag{4.115}$$

onde:

$\mathbf{Q}_{\hat{\mathbf{x}}} \in R^{t \times t}$: matriz cofatora das ambigüidades reais estimadas;

$\hat{\mathbf{x}} \in R^t$: vetor das ambigüidades reais estimadas;

$\mathbf{a} \in Z^t$: vetor de ambigüidades.

Os semi-eixos do elipsóide t-dimensional são obtidos como na equação (4.107). Os autovalores e autovetores da matriz $\mathbf{Q}_{\hat{\mathbf{x}}}$ podem ser calculados pela decomposição de valor singular (ver seção 3.1.3.3) equação (3.47).

A constante χ^2 pode ser calculada pela expressão (4.54) e, assim, o semi-eixo maior do elipsóide terá o comprimento:

$$\sigma_{MAX} = \chi \sqrt{\lambda_{MAX}}, \tag{4.116}$$

cuja direção é dado pelo autovetor (\mathbf{e}_{MAX}) de $\mathbf{Q}_{\hat{\mathbf{x}}}$ associado a λ_{MAX}

$$\mathbf{e}_{\text{MAX}} = \mathbf{V}(:, \mathbf{v}_{\text{MAX}}), \quad (4.117)$$

onde:

\mathbf{V} : é a matriz ortogonal com autovetores de \mathbf{Q}_i a partir da $\text{svd}(\mathbf{Q}_i)$.

Experiências mostram que nem sempre o mínimo da quadrática $f(\mathbf{a})$, pode ser encontrado pelos candidatos definidos pela equação (4.113). Todavia, experiências também comprovam que nem sempre o minimizador ($\tilde{\mathbf{a}}$) da quadrática $f(\mathbf{a})$, é o vetor de ambigüidades corretas (ver anexo 2)¹⁰. E, em certos casos os candidatos definidos pela equação (4.113), encontram não o mínimo da quadrática $f(\mathbf{a})$, mas sim, os valores corretos das ambigüidades conforme ilustra o exemplo a seguir.

Em continuidade ao exemplo do caso bi-dimensional; agora, serão consideradas as outras ambigüidades envolvidas. O vetor de ambigüidades reais estimados e a correspondente matriz cofatora foram:

$$\hat{\mathbf{x}} = \begin{bmatrix} -172710,649 \\ 115806,049 \\ -90961,765 \\ 3502437,418 \\ 220643,642 \\ 229997,659 \end{bmatrix}; \quad \mathbf{Q}_i = \begin{bmatrix} 1,610 & -0,957 & 3,008 & -1,964 & 1,046 & 1,146 \\ -0,957 & 0,611 & -1,838 & 1,213 & -0,681 & -0,742 \\ 3,008 & -1,838 & 5,692 & -3,695 & 2,051 & 2,241 \\ -1,964 & 1,213 & -3,695 & 2,516 & -1,277 & -1,400 \\ 1,046 & -0,681 & 2,051 & -1,277 & 0,821 & 0,890 \\ 1,146 & -0,742 & 2,241 & -1,400 & 0,890 & 0,966 \end{bmatrix}$$

¹⁰ Os anexos 1,2 e 3 correspondem aos extratos de resultados dos testes realizados para três linhas de base: linha de base 1 (ponto Norte 1 – ponto Sul 1) com comprimento de 5,002 m; linha de base 2 (ponto P003 – ponto Pc) com comprimento de 185,153 m; e linha de base 3 (ponto P003 – estação PARA/SIRGAS) com comprimento 950,565 m. Os extratos para cada linha de base são dois. O que aparece na frente é o resultado encontrado pelo programa que elaboramos e, o que aparece depois, com o objetivo apenas de comparação das ambigüidades e do vetor da linha de base, com os valores que encontramos em nosso programa, é o resultado fornecido do processamento com o software Prism da Ashtech. No extrato de nosso programa aparecem dois resultados para cada época. O primeiro resultado refere-se à solução flutuante do problema de mínimos quadrados e, o segundo, a solução com as ambigüidades injuncionadas aos inteiros (equação 4.64), obtidas da solução ($\tilde{\mathbf{a}}$) do problema de mínimos quadrados inteiro (equação 3.80, equação 4.48). Observe que o mínimo da quadrática do problema de mínimos quadrados inteiro (com os dados acumulados seqüencialmente), só passa a corresponder a solução correta para o teste da linha de base 1, na 19ª época; para o teste da linha de base 2, na 14ª época; e para o teste da linha de base 3, na 15ª época, quando então, passa a existir uma estabilização na solução das ambigüidades. Note também, que para o primeiro teste, o resultado da linha de base 1 para a 19ª época, só difere em 7.1 mm com o resultado da 73ª época. Isto quer dizer que não seria preciso rastrear 18 min e 15 s, bastaria ficar no ponto 4 min e 45 s para se ter um resultado com diferença menor que 1 cm. Para a linha de base 2, tem-se uma diferença de 8.7 mm com uma diferença de tempo de 28 min e 30 s. E, finalmente, para a linha de base 3, tem-se uma diferença de 5,9 mm com uma diferença de tempo de 24 min e 15 s.

O mínimo inteiro ($\tilde{\mathbf{a}}$) da quadrática $f(\mathbf{a})$ é alcançado em:

$$\tilde{\mathbf{a}} = \begin{bmatrix} -172709 \\ 115806 \\ -90960 \\ 3502436 \\ 220643 \\ 229997 \end{bmatrix}.$$

A direção ($\bar{\mathbf{d}}$) a partir de $\hat{\mathbf{x}}$ para $\tilde{\mathbf{a}}$, e a direção do semi-eixo maior do elipsóide 6-dimensional são, respectivamente:

$$\bar{\mathbf{d}} = \begin{bmatrix} 0,559 \\ -0,016 \\ 0,598 \\ -0,481 \\ -0,218 \\ -0,223 \end{bmatrix}; \quad \mathbf{e}_{\text{MAX}} = \begin{bmatrix} 0,364 \\ -0,224 \\ 0,691 \\ -0,451 \\ 0,250 \\ 0,273 \end{bmatrix}.$$

O ângulo (θ) entre os vetores $\bar{\mathbf{d}}$ e \mathbf{e}_{MAX} , é:

$$\theta = \arccos \left(\frac{\langle \bar{\mathbf{d}}, \mathbf{e}_{\text{MAX}} \rangle}{\|\bar{\mathbf{d}}\| \|\mathbf{e}_{\text{MAX}}\|} \right) \cong 43^\circ 44'.$$

As direções entre os dois vetores já são significativamente diferentes. No entanto, o candidato $\mathbf{a}_k = \text{round}(\hat{\mathbf{x}} + t_k \cdot \mathbf{e}_{\text{MAX}})$, com $t_k = -0,5$, tem os valores:

$$\mathbf{a}_k = \begin{bmatrix} -172711 \\ 115806 \\ -90962 \\ 3502438 \\ 220644 \\ 229998 \end{bmatrix}.$$

que é a solução correta das ambigüidades, conforme mostrado no anexo 2, a partir da 14ª época.

O exemplo acima, sugere tentar de alguma maneira verificar se os candidatos, $\mathbf{a}_k = \text{round}(\hat{\mathbf{x}} + t_k \cdot \mathbf{e}_{\text{MAX}})$, são as ambigüidades corretas. Para isso, é preciso associar uma verificação que seja independente da minimização da quadrática $f(\mathbf{a})$. Esta verificação é feita usando a função de ambigüidade ^{minimização} para a simples diferença de fase (ver seção 4.2.1, equação (4.26) ou (4.27)).

Com a equação (3.84) ou (3.99), pode-se usar o vetor candidato \mathbf{a}_k como injeção e, então, encontrar o vetor de coordenadas $(\hat{\mathbf{x}}_1)$ da estação incógnita. Com esse vetor $\hat{\mathbf{x}}_1 = \mathbf{x}_B$ calcula-se a simples diferença de distância ρ_{AB}^j , $j = 1, \dots, n$ (n = número de satélites) e o valor da função de ambigüidade:

$$f_{\text{amb}} = \sqrt{\left[\sum_{j=1}^n \cos(\varphi_{AB}^j - \frac{2\pi}{\lambda} \rho_{AB}^j) \right]^2 + \left[\sum_{j=1}^n \text{sen}(\varphi_{AB}^j - \frac{2\pi}{\lambda} \rho_{AB}^j) \right]^2}. \quad (4.118)$$

O vetor candidato \mathbf{a}_k com o correspondente vetor de coordenadas $\mathbf{x}_{B(k)}$ que fornecer o maior valor para a função $f_{\text{amb}(k)}$, se não for o vetor de ambigüidades correto, deve estar bem próximo dele. Usando $\mathbf{x}_{B(k)}$ como aproximação inicial, pode-se aplicar o algoritmo 3.11 para maximizar localmente a função (4.118) e encontrar $\tilde{\mathbf{x}}_B$ (lembrando mais uma vez que $\max f_{\text{amb}} = \min - f_{\text{amb}}$).

Com o vetor de coordenadas $\tilde{\mathbf{x}}_B$, pode-se proceder como no método da linha de base conhecida, calculando as duplas diferenças de ambigüidades:

$$N_{AB}^{l,i} = \text{round} \left[\frac{1}{\lambda} (\rho_{AB}^{l,i} - \Phi_{AB}^{l,i}) \right], i = 2, \dots, n. \quad (4.119)$$

Se ocorrer alguma diferença entre as componentes do candidato \mathbf{a}_k e $N_{AB}^{l,i}$, $i = 1, 2, \dots, n$, atualizar \mathbf{a}_k ($\mathbf{a}_k \leftarrow N_{AB}^{l,i}$) e obter novamente o valor de $\tilde{\mathbf{x}}_1$ por (3.84) ou (3.99).

Com o procedimento acima descrito, um método já fica definido. Mas, por outro lado, ainda se pode usar o conceito de satélites primários e satélites secundários (ver seção 4.3.5) para assegurar mais confiabilidade ao processo. Assim, com 4 satélites primários, realiza-se a busca direcional no elipsóide tri-axial das

ambigüidades, encontrando o candidato $\mathbf{a}_k = \text{round}(\hat{\mathbf{x}} + t_k \cdot \mathbf{e}_{\text{MAX}})$ e o correspondente vetor de coordenadas $\mathbf{x}_{B(k)}$, que fornece o maior valor para a função de ambigüidade $f_{\text{amb}(k)}$. Aplicar o algoritmo 3.11 para maximização local de $f_{\text{amb}(k)}$ e obter o vetor $\tilde{\mathbf{x}}_B$. Com o vetor $\tilde{\mathbf{x}}_B$ proceder como no método da linha de base conhecida, primeiro para as ambigüidades do grupo primário:

$$N_{AB}^{1,i} = \text{round} \left[\frac{1}{\lambda} (\rho_{AB}^{1,i} - \Phi_{AB}^{1,i}) \right], i = 2, 3, 4. \quad (4.120)$$

Se $N_{AB}^{1,i}$, $i = 2, 3, 4$ forem diferentes do vetor candidato \mathbf{a}_k , atualizar $\mathbf{a}_k = [N_{AB}^{1,2} \ N_{AB}^{1,3} \ N_{AB}^{1,4}]^T$ e calcular novamente $\tilde{\mathbf{x}}_B$.

Com $\tilde{\mathbf{x}}_B$, calcular as ambigüidades do grupo secundário:

$$N_{AB}^{1,i} = \text{round} \left[\frac{1}{\lambda} (\rho_{AB}^{1,i} - \Phi_{AB}^{1,i}) \right], i = 5, 6, \dots, n. \quad (4.121)$$

Injuncionando todas as ambigüidades dos grupos primários e secundários, calcula-se o vetor das coordenadas $\mathbf{x}_{B(c)}$ (ver equação 4.74). Se a diferença:

$$\|\mathbf{x}_{B(c)} - \tilde{\mathbf{x}}_B\|_2 < \varepsilon, \quad (4.122)$$

com:

$$\varepsilon \approx 10 \text{ cm},$$

então, verifica-se se as hipóteses $H_0: \tilde{\sigma}_0^2 = \sigma_0^2$ e $H_0: \tilde{\sigma}_0^2 = \hat{\sigma}_0^2$ não são rejeitadas a um nível de significância $\alpha=5\%$. Se ambas hipóteses, não forem rejeitadas a esse nível de significância e a solução ($\tilde{\mathbf{a}}_k$) da época atual (considerando o ajustamento seqüencial) é a mesma das épocas anteriores (suponha-se que não ocorreu perdas de ciclo), ou seja, houve estabilização, as ambigüidades são tidas como resolvidas.

Infelizmente nesse método, o teste estatístico da razão da variância mínima estimada a posteriori (ambigüidades injuncionadas) com e segunda menor variância estimada a posteriori (ambigüidades injuncionadas), não pode ser aplicado. Contudo, se a solução se já estabilizou durante as últimas três épocas, após não ter sido rejeitada as hipóteses (H_0) acima mencionadas, já se pode aceitar a resolução das ambigüidades.

5. CONCLUSÕES E RECOMENDAÇÕES

5.1 CONCLUSÕES

A pesquisa sobre resolução das ambigüidades GPS, como pôde ser verificado ao longo dessa pesquisa, em sua quase totalidade se depara com o ajustamento das observações GPS da fase da portadora. O ajustamento de observações, por sua vez, fundamentado por estudos e análises de dois ramos da Ciência. A Matemática Aplicada que trata dos assuntos relacionados aos métodos de otimização e aos métodos de resolução de sistemas de equações lineares. E, a Estatística que trata dos assuntos relacionados às leis probabilísticas das observações e a validade dos modelos envolvidos.

O problema posto fez surgir algumas indagações que nortearam a pesquisa para o contexto da Matemática Aplicada: i) O mal-condicionamento nos sistemas de equações normais afeta na solução do ajustamento das observações GPS da dupla diferença de fase? ii) No processo de otimização de uma função objetivo, o ponto inicial não sendo muito próximo da solução, às iterações alcançarão a convergência para o ponto ótimo? iii) Na resolução das ambigüidades é mesmo necessário realizar buscas em todo um espaço através de milhares de soluções potenciais do problema?

A resposta para cada uma dessas indagações estão totalmente fundamentadas no trabalho, a começar pela primeira, cujos estudos foram iniciados em GARNÉS (1996) e GARNÉS; SAMPAIO e DALMOLIN (1997). Lá ficou comprovado que em problemas de mínimos quadrados, mal-condicionados, o método de solução (algoritmo) desenvolve um papel primordial. Inclusive, em vários exemplos mostrados, o método de Cholesky que é o mais clássico para resolver o sistema de equações normais apresentou problemas graves de instabilidade. Nesse trabalho, no entanto, o estudo aplicado ao ajustamento GPS para a solução do problema de mínimos quadrados linear, das duplas diferenças de fase, abordou, dos métodos diretos, além do método da decomposição de Cholesky, também os métodos da decomposição QR; e da decomposição de valor singular- SVD, e estendeu-se aos métodos iterativos dos gradientes conjugados e gradientes conjugados com

pré-condicionamento.

As comparações entre tais métodos foram realizadas usando duas estratégias. Na primeira delas, aplicou-se, de fato, o ajustamento sequencial conforme apresentado na seção (3.1.4) e, na segunda, as observações foram colocadas todas juntas de forma a evitar a formação do sistema de equações normais. Depois da análise de vários testes envolvendo até as 20 primeiras épocas no ajustamento das linhas de base, anexos 1,2 e 3, em precisão computacional tipo real (11-12 dígitos significativos) e também tipo dupla (15-16 dígitos significativos), do Delphi Versão 4.0. Pôde-se verificar que mesmo com o número de condição elevado (até a ordem de 1.10^8) do sistema de equações normais, nenhum dos métodos apresentaram instabilidade na solução, isto é, todos encontravam dentro de um certo limite os mesmos valores. Resultado de certa forma até esperado, de acordo com análise teórica da seção (3.1.2). Neste ponto, chama-se a atenção para o número de iterações necessárias para o método dos gradientes conjugados alcançar a convergência. Em certos casos, quando a condição de parada para o método exigia $\|\mathbf{x}_k - \mathbf{x}_k\| < 1e^{-8}$, era preciso ir um pouco mais de uma vez e meia do limite teórico (n) de iterações pré-estabelecida para se ter a convergência, ao passo que, com o pré-condicionamento, o limite teórico era ultrapassado em no máximo uma iteração.

No caso do problema de mínimos quadrados não-linear, a comparação se deu entre os métodos de Gauss-Newton, Levenberg-Marquardt e Gauss-Newton modificado globalizado. A comparação foi realizada no sentido de verificar se ocorriam problemas em se utilizar como ponto de aproximação inicial do processo iterativo, um que não estivesse tão próximo da solução. Por surpresa! para o método de Gauss-Newton, chegou-se a conclusão que não é necessário o conhecimento aproximado nem para as coordenadas da estação incógnita, nem para as ambigüidades GPS da dupla diferença, fato que em outras aplicações da Geodésia, constitui graves prejuízos à convergência. Iniciando-se com o vetor nulo para os valores dos parâmetros, o método de Gauss-Newton sempre obteve sucesso alcançado a convergência para o ponto desejado. A convergência, só não ocorreu em pontos absurdamente afastados da solução (a solução para as coordenadas da estação incógnita, sendo: $x=3763761,227$ m; $y=-4365119,772$ m; $z=-2724382,923$ m), como por

$z = 0$ m. Nesses pontos, ambos os métodos, o de Levenberg-Marquardt, bem como, o de Gauss-Newton modificado Globalizado obtiveram sucesso. Nesta fase, implementou-se no método de Gauss-Newton, os métodos de solução do problema de mínimos quadrados linear, anteriormente comentados, já que a cada atualização do passo de Gauss-Newton deve-se resolver um subproblema deste tipo. Desta forma, quando foi usado como critério de parada as condições $\|\mathbf{x}_k - \mathbf{x}_{k-1}\| < 1e^{-8}$ e $\|\mathbf{x}_k - \mathbf{x}_{k-1}\| < 1e^{-9}$, no ajuste das duplas diferenças de fase, apenas os métodos iterativos: gradientes conjugados e gradientes conjugados com pré-condicionamento conseguiram atingir a estabilização e parar. Nos métodos diretos a estabilidade só foi atingida até a condição $\|\mathbf{x}_k - \mathbf{x}_{k-1}\| < 1e^{-8}$.

A proposta de realizar o ajustamento das observações GPS da dupla diferença de fase por meio da solução do problema de mínimos quadrados linear misto (seção 4.4.1.3), utilizando os métodos de penalidade com as funções barreiras (ver equação (4.79), junto com a função de penalidade sobre os não-inteiros (ver equação (4.87)), passaram por algumas tentativas que valem a pena ser relatadas. Na primeira tentativa procurou-se minimizar a função objetivo (4.100), usando o método de Newton clássico, mas tal tentativa muitas vezes foi frustrada, porque o passo de Newton acabava levando para fora da região factível e não mais voltava. Depois, implementou-se o método de Newton modificado globalizado com a decomposição de Cholesky modificada (algoritmo 3.11). Uma melhora considerável foi obtida, mas, mesmo assim, acabavam ocorrendo falhas, às vezes porque o valor da função objetivo após a barreira, tinha menor valor do que o ponto de onde o passo foi tentado. Utilizando a função de penalidade (4.81), muitas vezes a convergência era alcançada para um ponto sobre os inteiros, mas ainda assim, muitos problemas aconteciam por conta dos parâmetros utilizados. Logo, também esse tipo de função de penalidade foi abandonado. Passou-se, então, à utilização da função de penalidade hiperbólica na busca de resolver o problema dado pela equação (4.101). À primeira vista, os resultados foram muito promissores, pois todos os problemas encontrados com as outras funções de barreira e penalidade foram superados. O método de Newton modificado globalizado (algoritmo 3.11) sempre encontrava a solução para as ambigüidades sobre os inteiros (resultado desejado). Todavia, a convergência sempre

acontecera, mas nem sempre para o ponto de mínimo global sobre os inteiros. Como o objetivo estava em encontrar as ambigüidades que forneciam também a solução do problema de mínimos quadrados linear inteiro, não se obteve sucesso com esse método.

O método usando busca direcional nos espaços das ambigüidades e das coordenadas, conforme se discorreu na seção (4.4.2), surgiu quando da realização de buscas aos inteiros pela máxima direção de descida na função quadrática do problema de mínimos quadrados inteiro. O método usando a contribuição dos grupos de satélites primários e secundários, tem a garantia de que os vetores de ambigüidades candidatos não gerem nas coordenadas da estação incógnita, uma diferença maior do que 10 cm (ver equação 4.122), o que para muitas aplicações já seria satisfatório. Os 4 satélites do grupo primário são selecionados no início do processo (de modo que o GDOP fique entre 3 e 6, aproximadamente), quando então, realiza-se a verificação de qual satélite tem maior ângulo de elevação para ser usado como satélite de referência na formação das duplas diferenças de fase. O ajustamento seqüencial usado para obtenção da solução flutuante, permite extração da matriz variância-covariância dos satélites do grupo primário, e com a aplicação da decomposição de valor singular, obtém-se a direção e o comprimento da linha de busca (ver equação 4.117 e 4.116). Os vetores candidatos à busca direcional são aqueles definidos pela equação (4.113), incluindo, $t_k=0$, o qual corresponde, simplesmente, arredondar cada ambigüidade para seu valor inteiro. Os restantes dos detalhes estão bem relatados na descrição do método.

O método foi aplicado para as três linhas de base exemplificadas nos anexos. Em todas elas, o método alcançou sucesso e estabilizou-se à solução em menos de 5 minutos (a taxa de coleta dos dados foi de 15 s). Em algumas épocas a solução “correta” era encontrada antes mesmo da estabilização. Os testes de hipóteses usados (ver equação 4.122) foram bastante flexíveis e serviram apenas para confirmar as hipóteses gerais. A decisão final ficou mesmo por conta da estabilização.

5.2 RECOMENDAÇÕES

Apesar do método de Gauss-Newton ter tido um comportamento adequado quando se utilizou o vetor nulo como vetor de valores aproximados para a iniciação do processo iterativo, ele não tem características de convergência global. Neste sentido, recomenda-se na implementação dos programas de ajustamento GPS, o método de Gauss-Newton modificado globalizado, ou o método de Levenberg-Marquardt.

Dos métodos testados para a solução do problema de mínimos quadrados linear: Cholesky, QR, decomposição de valor singular, gradientes conjugados e gradientes conjugados com pré-condicionamento, no ajustamento de linhas de base GPS, o tempo computacional realmente não é importante, principalmente pela velocidade de processamento dos computadores atuais. O mesmo, por exemplo, já não pode ser dito no ajustamento simultâneo de uma rede GPS, cabendo investigações para verificar o mais adequado. No caso dessa pesquisa, uma atenção muito especial deve se dada a solução pela decomposição de valor singular, pois é por meio dela que fica definida a direção e o comprimento da linha de busca do método desenvolvido na seção (4.4.2). Ela pode, por exemplo, ser implementada para dar a solução da última iteração dos métodos de Gauss-Newton modificado globalizado, ou do método de Levenberg-Marquardt.

Para o problema de mínimos quadrados linear misto, não se teve a oportunidade durante o desenvolvimento desse trabalho, de realizar testes com outras funções de penalidades sobre os não-inteiros (ver equação 4.88). Recomendamos para futuras pesquisas, encontrar uma função mais adequada para que o método funcione e possa ser utilizado.

O método de busca direcional em espaços integrados é um método barato em relação a outros métodos que realizam buscas em todo o espaço, seja das ambigüidades, ou seja, das coordenadas. E é preciso ainda investigar muitas de suas potencialidades, como, por exemplo: seu comportamento em linhas de base médias e longas; sua viabilidade em aplicação de resolução de ambigüidades On-The-Fly; seu comportamento quando são inseridas observações das pseudodistâncias no ajustamento, etc.

REFERÊNCIAS BIBLIOGRÁFICAS

ABIDIN, H. Z. **Computational and geometrical aspects of on-the-fly ambiguity resolution**. Fredericton, 1992. 292 f. Ph.D dissertation (Doctor of Philosophy) – Department of Geodesy and Geomatics Engineering, University of New Brunswick.

ABIDIN, H. Z. On-the-fly ambiguity resolution: formulation and results. **Manuscripta Geodaetica**, v. 18, p. 380-405, 1993.

ANDRADE, J. B. de. **NAVSTAR-GPS**. Curitiba, 1988. 63 f. Apostila - Curso de Pós-Graduação em Ciências Geodésicas, Universidade Federal do Paraná.

ASHTech Z-12 Receiver operating manual. USA: Ashtech, Inc., 1994.

BAZARAA, M. S.; SHETTY, C. M. **Nonlinear Programming: theory and algorithms**. New York: John Wiley & Sons, Inc., 1979.

BEUTLER, G.; WEBER, R.; HUGENTOBLE, U.; ROTHACHER, M.; VERDUN, A. GPS satellite orbits. In: TEUNISSEN, P.J.G.; KLEUSBERG, A. **GPS for geodesy**. 2.ed. Berlin: Springer-Verlag, 1998, p.43-109.

BJÖRCK, A. **Least squares methods**. Sweden: Department of Mathematics, Linköping University, 23 may, 1991.

CAMARGO, P. de O. **Modelo regional da ionosfera para uso em posicionamento com receptores GPS de uma frequência**. Curitiba, 1999. 192 f. Tese (Doutorado em Geodésia) – Setor de Ciências da Terra, Universidade Federal do Paraná.

CARVALHO, L. D. **Análise das técnicas GPS atuais para os posicionamentos estáticos e cinemáticos em bases curtas**. Curitiba, 1999. 139 f. Dissertação (Mestrado em Geodésia) – Setor de Ciências da Terra, Universidade Federal do Paraná

CORBETT, S. J.; CROSS, P. A. GPS single epoch ambiguity resolution. **Survey Review**, v. 33, n. 257, p. 149-160, jul. 1995.

DALMOLIN, Q. **Ajustamento de observações pelo processo iterativo**. Curitiba, 1976. 96 f. Dissertação (Mestrado em Geodésia) – Setor de Tecnologia, Universidade Federal do Paraná.

DENNIS JR, J. E.; SCHNABEL, R. B. **Numerical methods for unconstrained optimization and nonlinear equations**. Englewood Cliffs : Prentice-Hall, 1983.

DIVIS, D. A. (Ed.). SA: Going the way of the dinosaur. **GPS World**. Washington, v.11, n.6, p. 16-19, jun. 2000.

EULER, H. J.; LANDAU, H. Fast GPS ambiguity resolution on-the-fly for real-time applications. **Proceedings Int. Geod. Symp. on Satellite Positioning**, 6. Columbus, Ohio, p. 650-659, 1992.

FREL, E.; BEUTLER, G. Rapid static positioning based on the fast ambiguity resolution approach "FARA": Theory and first results. **Manuscripta Geodaetica**, v.15, n. 6, p. 325-356, 1990.

FREL, E. **Rapid differential positioning with the global positioning system (GPS)**. Geodätisch-geophysikalische Arbeiten in der Schweiz. Schweizerischen Geodätischen Kommission, 1991.

FRIEDLANDER, A. **Elementos de programação não-linear**. Campinas: Editora da UNICAMP, 1994.

GARNÉS, S. J. dos A. **Ajustamento paramétrico por mínimos quadrados com análise na estabilidade da solução**. Curitiba, 1996. 148 f. Dissertação (Mestrado em Geodésia) – Setor de Tecnologia, Universidade Federal do Paraná.

GARNÉS, S. J. dos A.; SAMPAIO, R. J. B. de.; DALMOLIN, Q. Ajustamento paramétrico por mínimos quadrados com análise na estabilidade da solução. **Boletim de Ciências Geodésicas**, Curitiba, v.2, p.24-35, 1997.

GARNÉS, S. J. dos A.; DALMOLIN, Q. Trilateração em redes topográficas: solução do problema linear por ajustamento livre. **Cartografia e Cadastro**, Lisboa, v. 9, p 59-67, 1998.

GARNÉS, S. J. dos A. Sistema de projeção e orientação das plantas topográficas. In: Congresso Brasileiro de Cadastro Técnico Multifinalitário, 3., 1998(a), Florianópolis. **Relação de trabalhos**. Florianópolis: Universidade Federal de Santa Catarina, 1998. p. 1-29. 1 CD ROM.

GARNÉS, S. J. dos A. **Método dos mínimos quadrados totais em problemas lineares**. Curitiba, 1998(b). 69 f. Monografia (Disciplina TG-792 seminário II) – Setor de Ciências da Terra, Universidade Federal do Paraná.

GARNÉS, S. J. dos A. **Problema de mínimos quadrados não-linear misto MQ-MQI na determinação das ambigüidades GPS**. Curitiba, 1998 (c). 43 f. Seminário (Qualificação da Tese – Curso de Pós-Graduação em Ciências Geodésicas) – Setor de Ciências da Terra, Universidade Federal do Paraná.

GARNÉS, S. J. dos A.; FERREIRA, L. D. D.; **Cálculo da posição dos satélites GPS com dados no formato RINEX**. Trabalho apresentado no 8. Congresso Nacional de Engenharia de Agrimensura, Criciúma, 1999.

GARNÉS, S. J. dos A.; SAMPAIO, R. J. B. de. Métodos diretos e iterativos na solução de subproblemas mal-condicionados de mínimos quadrados no ajustamento das DDF GPS. In: Simpósio Brasileiro de Pesquisa Operacional, 31., 1999, Juiz de Fora. **Relação de trabalhos**. Juiz de Fora: SOBRAPO, 1999. p. 215-227. 1 CD-ROM.

GASTINEL, N. **Linear numerical analysis**. New York: Academic Press, 1971, 350 p.

GEMAEL, C. **“Introdução à geodésia geométrica”: 1ª parte**. Curitiba, 1977. 148 f. Apostila – Curso de Pós-Graduação em Ciências Geodésicas, Universidade Federal do Paraná.

GEMAEL, C. **Geodesia celeste: Introdução**. Curitiba, 1991. 264 f. Apostila – Curso de Pós-Graduação em Ciências Geodésicas, Universidade Federal do Paraná.

GEMAEL, C. **Introdução ao ajustamento de observações: aplicações geodésicas**. Curitiba: Editora UFPR, 1994.

GILL, P. E.; MURRAY, W.; WRIGHT, M. H. **Practical optimization**. London: Academic Press, 1981.

GILL, P. E.; MURRAY, W.; WRIGHT, M. H. **Numerical linear algebra and optimization**. v. 1. Redwood: Addison-Wesley Publishing Company, 1991. 426 p.

GOLUB, G.H.; REINCH, C. Singular value decomposition and least squares solutions. **Handbook series linear Algebra**, New York, Numer. Math. 14, 403-20, 1970.

GOLUB, G. H.; VAN LOAN, C. F. **Matriz computations**. 2. ed. Baltimore: Johns Hopkins Press, 1989.

GPS WORLD: News and applications of the global positioning system. Washington: Advantar Communications, v. 11, n.6, jun. 2000.

GPS WORLD SHOWCASE 2001. Washington: Advanstar Communications, v. 11, n. 12, dez. 2000.

GURTNER, W.; MADER, G.; MACARTHUR, D. A common exchange format for GPS data. In: Int. Geod. Symp. Satellite Positioning, 5 th., 1989. **Proc...** Las Cruces, USA, v. 2, p. 920-931.

GURTNER, W.; MADER, G. Receiver independent exchange format version 2. **CSTG GPS Bulletin**. National Geodetic Survey, Rockville, v.3, n. 3, set/oct 1990.

HAMILTON, W. C. **Statistics in physical science: estimation, hypothesis testing, and least squares**. New York: the Ronald Press Company, 1964.

HAN, S.; RIZOS, C. Comparação das técnicas de resolução das ambigüidades GPS. **Geoconvergência**, out. 1998.

HATCH, R. Instantaneous ambiguity resolution. **IAG Symposium**, n. 107 “Kinematic Systems in Geodesy, Surveying, and Remote Sensing”, Banff, Canada, set. 10-13, (Kis’90), Springer Verlag, p. 299-308, 1990.

HOFMANN-WELLENHOF, B.; LICHTENEGGER, H.; COLLINS, J. **Global positioning system: theory and practice**. 3. ed. New York: Springer-Verlag Wien, 1994. 355 p.

IBGE. **SIRGAS relatório final: grupos de trabalho I e II**. Rio de Janeiro, 1997.

IGS Products. Disponível em: <<http://igsceb.jpl.nasa.gov/components/prods.html>> Acesso em: 1 jun. 2001.

JONGE, P. de; TIBERIUS, C. The LAMBDA method for integer ambiguity estimation: implementation aspects. **Delft Geodetic Computing Centre LGR series**, n.12, aug. 1996. 49 p.

JOHNSON, R.A.; WICHERN, D.N. **Applied multivariate statistical analysis**. 2. ed. New Jersey: Prentice-Hall, Inc., 1988.

KRUEGER, C. P. **Posicionamento cinemático de trens**. Curitiba, 1994. 168 f. Dissertação (Mestrado em Geodésia) – Setor de Tecnologia, Universidade Federal do Paraná.

KRUEGER, C. P. **Investigações sobre aplicações de alta precisão do GPS no âmbito marinho**. Curitiba, 1996. 267 f. Tese (Doutorado em Geodésia) – Setor de Ciências da Terra, Universidade Federal do Paraná.

LANGLEY, R. B. Propagation of the GPS signals. In: TEUNISSEN, P.J.G.; KLEUSBERG, A. **GPS for geodesy**. 2.ed. Berlin: Springer-Verlag, 1998, p.111-149.

LAWSON, C. L. ; HANSON, R. J. **Solving least squares problems**. Englewood Cliffs: Prentice-Hall, Inc., 1974.

LEICK, A. **GPS satellite surveying**. 2. ed. New York: John Wiley & Sons, Inc., 1995.

LUENBERGER, D. G. **Introduction to linear and nonlinear programming**. Menlo Park: Addison-Wesley Publishing Company, Inc., 1973.

LUGNANI, J.B. **O problema dos sistemas de equações lineares mal condicionados e suas implicações em geodésia**. Curitiba, 1975. 106 f. Dissertação (Mestrado em Geodésia) – Setor de Tecnologia, Universidade Federal do Paraná.

MACHADO, W.C.; MONICO, J.F.G.; CAMARGO, P. de O. SA: Antes e depois de sua desativação. **A Mira**, Criciúma, v.98, p.38-43, agosto/setembro.2000.

MADER, G. L. GPS antenna calibration at the National Geodetic Survey. **National Geodetic Survey, NOS, NOAA**. Silver Spring. MD. Disponível em: <<http://www.ngs.noaa.gov/ANTCAL/Files/summary.html>> Acesso em: 1 jun. 2001.

MADER, G.L. **GPS Antena Calibration**. Disponível em: < http://www.ngs.noaa.gov/ANTCAL/Models/TRM29659_00.shtml > Acesso em: 1 jun. 2001.

MADER, G. L. GPS antenna calibration. **National Geodetic Survey, NOS, NOAA**. Disponível em: < http://www.ngs.noaa.gov/ANTCAL/Models/Diagrams/TRM29659_00.gif > Acesso em: 1 jun. 2001.

MARTINEZ, J. M.; SANTOS, S. A. **Métodos computacionais de otimização**. In: Colóquio Brasileiro de Matemática, 20, 1995, Rio de Janeiro: Instituto de Matemática Pura e Aplicada. 24-28 jul., 1995. 256 p.

MIKHAIL, E. M.; ACKERMAN, F. **Observations and least squares**. New York: University Press of America, Inc., 1976.

MIKHAIL, E. M.; GRACIE, G. **Análisis and adjustment of survey measurements**. New York: Van Nostrand Reinhold Company, 1981.

MONICO, J. F. G. **High precision inter-continental GPS network**. Nottingham, 1995. 205 f. Thesis (Doctor Philosophy) – Institute of Engineering Surveying and Space Geodesy, University of Nottingham.

MONICO, J. F. G. Métodos para a solução da ambigüidade: primeiras experiências na UNESP de Presidente Prudente. In: GIS Brasil 96, 1996, Curitiba. **Anais...**, 1996. p. 465-474.

NOBLE, B. ; DANIEL, J. W. **Applied linear algebra**. 2. ed. Englewood Cliffs: Prentice-Hall, Inc., 1977.

PRESS, W. H.; FLANNERY, B. P.; TEUKOLSKY, S. A.; VETTERLING, W. T. **Numerical recipes: The art of scientific computing**. Cambridge: Cambridge University Press, 1986.

RAO, S. S. **Optimization: theory and applications**. New York: John Wiley & Sons, 1979.

REMONDI, B. W. Performing centimeter-level surveys in seconds with GPS carrier phase: initial results. **Navigation: Journal of the Institute of Navigation**. v. 32, n. 4, 1985-1986. p. 386-400.

RIZOS, C.; HAN, S. A. A new method for constructing multi-satellite ambiguity combinations for improved ambiguity resolution. **Proceedings of ION GPS-95**. The 8th International Technical Meeting of the Satellite Division of the Institute of Navigation, Parte 2 of 2. Palm Spring Convention Center, Palm Springs, California, p. 12-15, set. 1995.

ROTHACHER, M.; MERVART, L.; BEUTLER, G.; BROCKMANN, E.; FANKHAUSER, S.; GURTNER, W.; JOHNSON, J.; MERVART, L.; ROTHACHER, M.; SCHAER, S.; SPRINGER, T.; WEBER, R. **Bernese GPS software: version 4.0**. Berne: Astronomical Institute University of Berne, September, 1996.

SANTOS, M. C. dos. **NAVSTAR/GPS: aspectos teóricos e aplicações geofísicas**. Publicação Especial n. 7, observatório Nacional do Rio de Janeiro, 1990. 127 p.

SEEBER, G. **Satellite geodesy: foundations, methods, and applications**. Berlin: Walter de Gruyter, 1993.

SEEBER, G. Posicionamento e navegação de precisão (2000). In.: EVENTO DE EXTENSÃO UNIVERSITÁRIA, 2000, sob o auspícios do Institut für Erdmessung da Universidade de Hannover e do Departamento de Geomática do Setor de Ciências da Terra da Universidade Federal do Paraná, Curitiba: Curso de Pós-Graduação em Ciências Geodésicas – UFPR, 13 – 17 de mar. 2000, **Transparência** sn.

SCHNABEL, R. B.; ESKOW, E. A revised modified cholesky factorization algorithm. **SIAM J. Optim.**, v. 9, n. 4, p. 1135-1148, 1999.

SILVA, B.P.de B.e. **A preconditioner for least squares problems**. Curitiba, 1997. 56 f. Dissertação (Mestrado em Métodos Numéricos em Engenharia) – Setor de Ciências Exatas e de Tecnologia, Universidade Federal do Paraná.

SPIPKER JUNIOR, J. J. GPS signal structure and performance characteristics. **Navigation**, Alexandria. v. 1.,p 29-54, 1980.

STRANG, G. **Linear algebra and its applications**. Orlando: Academic Press, INC., 1980. 414 p.

STRANG, G.; BORRE, K. **Linear algebra, geodesy, and GPS**. Wellesley: Cambridge Press, 1997.

TEUNISSEN, P. J. G. The invertible GPS ambiguity transformations. **Manuscripta Geodaetica**. v. 20, p. 489-497, 1995.

TEUNISSEN, P. J. G.; JONGE, P. J. de; TIBERIUS, C. C. J. M. A new way to fix carrier-phase ambiguities. **GPS WORLD**, v. 6, n. 4, p. 58-61, abr. 1995.

TEUNISSEN, P. J. G.; KLEUSBERG, A. **GPS for geodesy**. 2. ed. Berlin: Springer-Verlag Belin Heidelberg, 1998. 650 p.

TIERRA CRIOLLO, A. R. **Comparação da precisão dos métodos cinemáticos para distâncias menores a dez quilômetros**. Curitiba, 1993. 120 f. Dissertação (Mestrado em Geodésia) – Setor de Tecnologia, Universidade Federal do Paraná.

UNITED STATES COAST GUARD NAVIGATION CENTER. Disponível em: <<http://www.navcen.uscg.gov/gps/precise/default.html>> Acesso em: 1 ago. 2001.

WELLS, D.; BECK, N.; DELIKARAOGLU, D.; KLEUSBERG, A.; KRAKIWSKY, E. J. ; LACHAPPELLE, G.; LANGLEY, R. B.; NAKIBOGLU, M. SCHWARZ, K. P.; TRANQUILLA, J. M.; VANÍCEK, P. **Guide to GPS positioning**. Canadá: Canadian GPS Associates, Fredericton, N. B., 1986.

XAVIER, A. E. **hyperbolic penalty: a new method for nonlinear programming with inequalities**. Trabalho concedido pelo autor, 1998 (no prelo).

YUAN, J.Y. **Iterative methods for the generalized least-squares problem**. Rio de Janeiro, 1993. 139 f. Tese (Doutorado em Matemática Aplicada) – Instituto de Matemática Pura e Aplicada.

ANEXOS

ANEXO 1 – LINHA DE BASE (5,002 m): EXTRATO DE RESULTADOS	199
ANEXO 2 – LINHA DE BASE (185,153 m): EXTRATO DE RESULTADOS	201
ANEXO 3 – LINHA DE BASE (950,565 m): EXTRATO DE RESULTADOS	203

ANEXO 1 – LINHA DE BASE (5,002 m): EXTRATO DE RESULTADOS

Linha de Base 1 (Norte 1 – Sul 1): comprimento 5,002 m

Extrato dos resultados - solução flutuante e solução com ambigüidades injuncionadas aos inteiros, obtida da solução do problema de mínimos quadrados linear inteiro.

L i n h a d e B a s e				Duplas diferenças de ambigüidades (PRN ref=23)						
Épc.	dx	dy	dZ	PRN 21	PRN 30	PRN 22	PRN 17	PRN 6	PRN 3	PRN 26 PRN 29
2	-2.686	2.840	-4.360	-41400.90	92006.59	-938.66	52012.31	37141.73	708432.69	280434.96
2	0.169	5.266	-3.187	-41396.00	91997.00	-922.00	52013.00	37127.00	708450.00	280419.00
3	-1.280	1.906	-4.484	-41398.42	92009.29	-932.20	52016.41	37142.68	708443.56	280440.32
3	0.981	0.586	-5.413	-41392.00	92017.00	-923.00	52020.00	37144.00	708459.00	280447.00
4	-1.042	1.907	-4.664	-41397.36	92010.26	-931.33	52016.10	37142.35	708444.74	280440.03
4	1.478	1.947	-5.296	-41390.00	92014.00	-920.00	52017.00	37138.00	708460.00	280438.00
5	-1.552	2.003	-4.369	-41399.26	92008.47	-933.32	52016.17	37142.74	708441.83	280439.92
5	1.479	1.948	-5.294	-41390.00	92014.00	-920.00	52017.00	37138.00	708460.00	280438.00
6	-1.297	1.876	-4.393	-41398.77	92008.89	-932.15	52016.78	37142.75	708443.72	280440.61
6	1.479	1.949	-5.292	-41390.00	92014.00	-920.00	52017.00	37138.00	708460.00	280438.00
7	-1.360	1.889	-4.392	-41398.89	92008.85	-932.45	52016.66	37142.82	708443.27	280440.54
7	1.480	1.950	-5.291	-41390.00	92014.00	-920.00	52017.00	37138.00	708460.00	280438.00
8	-2.133	2.749	-3.838	-41401.39	92004.27	-935.09	52015.06	37140.73	708437.79	280435.82
8	-4.908	3.059	-2.857	-41410.00	91998.00	-947.00	52014.00	37144.00	708421.00	280436.00
9	-2.461	3.137	-3.605	-41402.42	92002.29	-936.21	52014.31	37139.75	708435.42	280433.66
9	-4.908	3.058	-2.857	-41410.00	91998.00	-947.00	52014.00	37144.00	708421.00	280436.00
10	-2.510	3.305	-3.516	-41402.63	92001.52	-936.26	52014.06	37139.17	708434.98	280432.67
10	-4.340	3.395	-3.351	-41407.00	92000.00	-945.00	52012.00	37142.00	708423.00	280433.00
11	-2.492	3.334	-3.512	-41402.58	92001.45	-936.15	52014.01	37139.02	708435.06	280432.48
11	-2.251	5.112	-3.264	-41401.00	91997.00	-934.00	52010.00	37132.00	708434.00	280421.00
12	-2.529	3.323	-3.499	-41402.70	92001.40	-936.31	52014.03	37139.12	708434.87	280432.57
12	-1.594	4.061	-3.824	-41399.00	92002.00	-932.00	52012.00	37135.00	708439.00	280427.00
13	-2.510	3.232	-3.571	-41402.53	92001.94	-936.37	52014.08	37139.47	708434.99	280433.09
13	-2.085	2.700	-3.942	-41401.00	92005.00	-935.00	52015.00	37141.00	708438.00	280436.00
14	-2.473	3.110	-3.633	-41402.39	92002.49	-936.32	52014.27	37139.89	708435.32	280433.81
14	-2.086	2.701	-3.941	-41401.00	92005.00	-935.00	52015.00	37141.00	708438.00	280436.00
15	-2.196	2.717	-3.888	-41401.41	92004.60	-935.49	52014.91	37141.00	708437.31	280436.00
15	-1.519	3.040	-4.433	-41398.00	92007.00	-933.00	52013.00	37139.00	708440.00	280433.00
16	-2.055	2.516	-4.009	-41400.94	92005.63	-935.06	52015.27	37141.56	708438.34	280437.13
16	-2.087	2.700	-3.941	-41401.00	92005.00	-935.00	52015.00	37141.00	708438.00	280436.00
17	-1.995	2.406	-4.076	-41400.72	92006.19	-934.91	52015.44	37141.90	708438.80	280437.76
17	-2.088	2.700	-3.942	-41401.00	92005.00	-935.00	52015.00	37141.00	708438.00	280436.00
18	-1.944	2.335	-4.120	-41400.54	92006.57	-934.75	52015.56	37142.10	708439.17	280438.16
18	-2.089	2.700	-3.942	-41401.00	92005.00	-935.00	52015.00	37141.00	708438.00	280436.00
19	-1.902	2.267	-4.166	-41400.38	92006.94	-934.64	52015.66	37142.31	708439.47	280438.55
19	-1.426	1.655	-4.499	-41399.00	92010.00	-933.00	52017.00	37144.00	708443.00	280442.00
20	-1.866	2.209	-4.199	-41400.27	92007.22	-934.53	52015.77	37142.48	708439.74	280438.88
20	-1.426	1.655	-4.499	-41399.00	92010.00	-933.00	52017.00	37144.00	708443.00	280442.00
.
.
.
70	-1.520	1.806	-4.446	-41399.18	92009.37	-933.36	52016.57	37143.52	708442.23	280441.14 1696688.31
70	-1.425	1.662	-4.498	-41399.00	92010.00	-933.00	52017.00	37144.00	708443.00	280442.00 1696689.00
71	-1.521	1.808	-4.445	-41399.18	92009.36	-933.36	52016.57	37143.51	708442.23	280441.13 1696688.30
71	-1.425	1.662	-4.498	-41399.00	92010.00	-933.00	52017.00	37144.00	708443.00	280442.00 1696689.00
72	-1.522	1.809	-4.444	-41399.19	92009.35	-933.36	52016.57	37143.51	708442.22	280441.13 1696688.28
72	-1.425	1.662	-4.498	-41399.00	92010.00	-933.00	52017.00	37144.00	708443.00	280442.00 1696689.00
73	-1.522	1.809	-4.442	-41399.19	92009.34	-933.36	52016.57	37143.51	708442.22	280441.13 1696688.28
73	-1.425	1.662	-4.498	-41399.00	92010.00	-933.00	52017.00	37144.00	708443.00	280442.00 1696689.00

- Obs. : 1) cada época apresenta: primeiro a solução flutuante (real) e, segundo a solução com as ambigüidades injuncionadas aos inteiros;
 2) os valores inteiros das ambigüidades só se estabilizaram na 19ª época (4 minutos e 45 segundos), a partir daí, continuaram estáveis até o final do rastreo;
 3) os valores obtidos pelo programa elaborado, para essa linha de base, foram idênticos aos obtidos pelo *software* Prism da Ashtech, veja a página seguinte.

Linha de Base 1 (Norte 1 – Sul 1): comprimento 5,002 m

Estação fixa: **Norte 1**

$\varphi = 25^{\circ} 26' 53,48585''$ S;

$\lambda = 49^{\circ} 13' 51.34261''$ W

H = 926,2497 m

Extrato dos resultados do processamento no *software Prism*

```

THE FIXED DOUBLE DIFFERENCE SOLUTION (L1)
Measure of geometry: 0.008864      Wavelength = 0.190294 (m/cycle)
num_meas = 525      num_used = 525      rms_resid = 0.008779(m)
Post-Fit Chisq = 1614.673      NDF      = 3.646

Reference SV: 23      Integer Search Ratio = 100.000
SV      Ambiguity      FIT      Meas      SV      Ambiguity      FIT      Meas
3      -708443.000X      0.035      73      6      -37144.000X      0.068      73
17      -52017.000X      0.018      73      21      41399.000X      0.031      73
22      933.000X      0.034      73      26      -280442.000X      0.049      73
29      -1696689.000X      0.040      14      30      -92010.000X      0.066      73

Sigmax (m):      0.008378
Sigmay (m):      0.008732
Sigmaz (m):      0.007021
x      y      z
x 1.00
y-0.59y 1.00
z-0.54z 0.47z 1.00

del_station: 0.000299 -0.000064 -0.000119
      Station1: FIXED STATION      Station2: UNKNOWN STATION
              (00000)      (NTE1)      (00000)      (SUL1)
Latitude: -25.44819051 -25 26 53.48585      -25.44823566 -25 26 53.64837
E-Long : 310.76907150 310 46 8.65739      310.76907156 310 46 8.65760
W-Long : 49.23092850 49 13 51.34261      49.23092844 49 13 51.34240
E-Height: 924.7554      924.7116

Baseline vector:      -1.4248      1.6616      -4.4977

Mark1_xyz : 3763758.5965 -4365117.8050 -2724386.4467
Az1 El1 D1 : 179.93177 -0.5022 5.0020
E1 N1 U1 : 0.0060 -5.0018 -0.0438
Mark2_xyz : 3763757.1717 -4365116.1434 -2724390.9444
Az2 El2 D2 : 359.93177 0.5022 5.0020
E2 N2 U2 : -0.0060 5.0018 0.0438
Tue May 29 23:16:08 2001

```

Obs.: Não usar as coordenadas acima para outros trabalhos, porque no processamento não estão sendo consideradas as alturas das antenas. O intuito é de apenas mostrar as duplas diferenças de ambigüidades, e as componentes da linha de base.

ANEXO 2 – LINHA DE BASE (185,153 m): EXTRATO DE RESULTADOS

Linha de Base 2 (P003 – PC) : comprimento 185,152 6 m

Extrato dos resultados - solução flutuante e solução com ambigüidades injuncionadas aos inteiros, obtida da solução do problema de mínimos quadrados linear inteiro.

L i n h a d e B a s e				Duplas diferenças de ambigüidades (PRN ref = 5)						
Épc.	dX	dY	dZ	PRN 30	PRN 24	PRN 6	PRN 10	PRN 8	PRN 9	PRN 17
2	-125.830	-138.059	29.496	-172720.58	115812.25	-90980.70	3502449.90	220636.80	229990.17	
2	-126.993	-140.395	29.162	-172727.00	115816.00	-90992.00	3502459.00	220634.00	229987.00	
3	-126.740	-139.878	29.118	-172725.72	115815.12	-90989.38	3502457.60	220635.21	229988.32	
3	-126.993	-140.394	29.162	-172727.00	115816.00	-90992.00	3502459.00	220634.00	229987.00	
4	-121.067	-128.634	28.645	-172696.68	115796.53	-90931.79	3502424.51	220659.47	230014.57	
4	-118.610	-123.480	28.345	-172684.00	115788.00	-90906.00	3502410.00	220671.00	230027.00	
5	-122.951	-132.202	29.239	-172705.75	115802.58	-90951.06	3502432.83	220649.47	230003.87	
5	-126.993	-140.396	29.158	-172727.00	115816.00	-90992.00	3502459.00	220634.00	229987.00	
6	-123.682	-134.121	29.513	-172709.60	115805.76	-90959.97	3502437.12	220644.37	229998.46	
6	-124.632	-133.550	29.357	-172713.00	115805.00	-90963.00	3502438.00	220646.00	230000.00	
7	-122.737	-131.396	30.270	-172703.29	115801.61	-90949.26	3502425.24	220645.67	230000.08	
7	-120.335	-129.411	30.046	-172693.00	115798.00	-90932.00	3502418.00	220651.00	230006.00	
8	-123.590	-133.492	29.880	-172708.43	115804.89	-90958.36	3502433.57	220643.66	229997.78	
8	-124.630	-133.557	29.356	-172713.00	115805.00	-90963.00	3502438.00	220646.00	230000.00	
9	-123.775	-133.881	29.731	-172709.56	115805.48	-90960.09	3502435.55	220643.68	229997.74	
9	-124.630	-133.557	29.356	-172713.00	115805.00	-90963.00	3502438.00	220646.00	230000.00	
10	-123.921	-134.158	29.666	-172710.38	115805.91	-90961.44	3502436.77	220643.47	229997.50	
10	-124.630	-133.557	29.356	-172713.00	115805.00	-90963.00	3502438.00	220646.00	230000.00	
11	-123.901	-134.114	29.649	-172710.29	115805.83	-90961.20	3502436.72	220643.65	229997.68	
11	-124.630	-133.557	29.355	-172713.00	115805.00	-90963.00	3502438.00	220646.00	230000.00	
12	-123.953	-134.258	29.596	-172710.65	115806.05	-90961.76	3502437.42	220643.64	229997.66	
12	-123.566	-134.240	29.750	-172709.00	115806.00	-90960.00	3502436.00	220643.00	229997.00	
13	-123.859	-134.074	29.581	-172710.18	115805.74	-90960.81	3502436.91	220644.08	229998.13	
13	-123.566	-134.240	29.750	-172709.00	115806.00	-90960.00	3502436.00	220643.00	229997.00	
14	-123.919	-134.243	29.550	-172710.56	115806.00	-90961.50	3502437.56	220643.91	229997.94	
14	-124.002	-134.288	29.506	-172711.00	115806.00	-90962.00	3502438.00	220644.00	229998.00	
15	-124.034	-134.437	29.527	-172711.16	115806.32	-90962.54	3502438.32	220643.65	229997.64	
15	-124.003	-134.287	29.506	-172711.00	115806.00	-90962.00	3502438.00	220644.00	229998.00	
16	-124.063	-134.480	29.515	-172711.31	115806.39	-90962.78	3502438.53	220643.63	229997.61	
16	-124.003	-134.287	29.506	-172711.00	115806.00	-90962.00	3502438.00	220644.00	229998.00	
17	-124.102	-134.544	29.503	-172711.52	115806.49	-90963.12	3502438.80	220643.56	229997.54	
17	-124.003	-134.287	29.506	-172711.00	115806.00	-90962.00	3502438.00	220644.00	229998.00	
18	-124.141	-134.584	29.499	-172711.70	115806.56	-90963.41	3502438.98	220643.50	229997.46	
18	-124.003	-134.286	29.506	-172711.00	115806.00	-90962.00	3502438.00	220644.00	229998.00	
19	-124.247	-134.788	29.486	-172712.26	115806.88	-90964.45	3502439.69	220643.17	229997.10	
19	-124.003	-134.286	29.507	-172711.00	115806.00	-90962.00	3502438.00	220644.00	229998.00	
20	-124.287	-134.877	29.479	-172712.49	115807.03	-90964.86	3502440.00	220643.04	229996.95	
20	-124.003	-134.285	29.507	-172711.00	115806.00	-90962.00	3502438.00	220644.00	229998.00	
.
.
.
125	-124.041	-134.325	29.515	-172711.09	115806.12	-90962.23	3502438.16	220644.01	229997.97	4296820.75
125	-124.000	-134.295	29.512	-172711.00	115806.00	-90962.00	3502438.00	220644.00	229998.00	4296821.00
126	-124.038	-134.316	29.516	-172711.07	115806.11	-90962.20	3502438.13	220644.02	229997.98	4296820.77
126	-124.000	-134.294	29.512	-172711.00	115806.00	-90962.00	3502438.00	220644.00	229998.00	4296821.00
127	-124.037	-134.307	29.517	-172711.06	115806.09	-90962.17	3502438.10	220644.03	229997.99	4296820.78
127	-124.000	-134.294	29.512	-172711.00	115806.00	-90962.00	3502438.00	220644.00	229998.00	4296821.00
128	-124.034	-134.297	29.518	-172711.04	115806.07	-90962.13	3502438.07	220644.04	229998.00	4296820.81
128	-124.000	-134.294	29.512	-172711.00	115806.00	-90962.00	3502438.00	220644.00	229998.00	4296821.00

Obs. : 1) cada época apresenta em primeiro a solução flutuante (real) e, em segundo a solução com as ambigüidades injuncionadas aos inteiros;
 2) os valores inteiros das ambigüidades só se estabilizaram na 14ª época (3 minutos e 30 segundos);
 3) os valores obtidos pelo programa elaborado, para essa linha de base, resultaram em uma diferença de 4,4 mm em relação aos obtidos pelo *software* Prism da Ashtech, veja a página seguinte.

Linha de Base 2 (P003 – PC) : comprimento 185,152 6 m

Estação fixa: P003

$\phi = 25^{\circ} 27' 19,35874''$ S

$\lambda = 49^{\circ} 14' 11,05371''$ W

H = 912,357 4 m

Extrato dos resultados do processamento no *software Prism*

```

THE FIXED DOUBLE DIFFERENCE SOLUTION (L1)
Measure of geometry: 0.012272      Wavelength = 0.190294 (m/cycle)
num_meas = 835      num_used = 806      rms_resid = 0.014156(m)
Post-Fit Chisq = 5944.998      NDF = 5.597

Reference SV: 5      Integer Search Ratio = 99.823
SV      Ambiguity      FIT      Meas      SV      Ambiguity      FIT      Meas
6      90962.000X      0.044      128      8      -220644.000X      0.115      100
9      -229998.000X      0.107      127      10      -3502438.000X      0.053      128
17     -4296821.000X      0.097      67      24      -115806.000X      0.046      128
30     172711.000X      0.039      128

Sigma_x (m):      0.018243
Sigma_y (m):      0.021887
Sigma_z (m):      0.009373
x      y      z
x 1.00
y-0.83y 1.00
z-0.64z 0.63z 1.00

del_station: -0.002350 0.003285 -0.000422

      Station1: FIXED STATION      Station2: UNKNOWN STATION
              (00000)      (P003)              (00000)      (PAUX)
Latitude: -25.45537743 -25 27 19.35874      -25.45505643 -25 27 18.20316
E-Long : 310.76359619 310 45 48.94629      310.76179079 310 45 42.44685
W-Long : 49.23640381 49 14 11.05371      49.23820921 49 14 17.55315
E-Height: 912.3574      918.4086

Baseline vector:      -124.0026      -134.2905      29.5113

Mark1_xyz : 3763110.6927 -4365209.7915 -2725100.1167
Az1 El1 D1 :      281.07999      1.8721      185.1526
E1 N1 U1 :      -181.6037      35.5647      6.0512
Mark2_xyz : 3762986.6902 -4365344.0820 -2725070.6054
Az2 El2 D2 :      101.08076      -1.8737      185.1526
E2 N2 U2 :      181.6043      -35.5647      -6.0512
Tue May 29 21:21:29 2001

```

Obs.: Não usar as coordenadas acima para outros trabalhos, porque no processamento não estão sendo consideradas as alturas das antenas. O intuito é de apenas mostrar as duplas diferenças de ambigüidades, e as componentes da linha de base.

ANEXO 3 – LINHA DE BASE (950,565 m): EXTRATO DE RESULTADOS

Linha de Base 3 (P003 – PARA) : comprimento 950,564 5 m

Extrato dos resultados - solução flutuante e solução com ambigüidades injuncionadas aos inteiros, obtida da solução do problema de mínimos quadrados linear inteiro.

L i n h a d e B a s e			Duplas diferenças de ambigüidades (PRN ref = 5)							
Épc.	dX	dY	dZ	PRN 30	PRN 24	PRN 6	PRN 10	PRN 8	PRN 9	PRN 17
2	641.567	98.241	695.719	-13653689.80	1452135.58	-12875586.80	-17804564.32	2281192.32	4445197.61	
2	638.024	91.492	696.727	-13653707.00	1452147.00	-12875623.00	-17804548.00	2281174.00	4445178.00	
3	639.342	93.578	695.307	-13653701.99	1452143.08	-12875609.25	-17804547.62	2281185.42	4445189.97	
3	638.009	89.854	694.983	-13653710.00	1452149.00	-12875625.00	-17804535.00	2281180.00	4445184.00	
4	640.144	94.186	695.432	-13653698.37	1452142.02	-12875603.91	-17804550.96	2281186.05	4445190.77	
4	638.011	89.852	694.982	-13653710.00	1452149.00	-12875625.00	-17804535.00	2281180.00	4445184.00	
5	640.754	95.702	695.388	-13653695.02	1452139.55	-12875596.92	-17804555.28	2281189.19	4445194.16	
5	641.240	94.682	695.275	-13653694.00	1452141.00	-12875597.00	-17804553.00	2281188.00	4445193.00	
6	640.882	95.649	695.414	-13653694.55	1452139.61	-12875596.47	-17804555.49	2281188.98	4445193.96	
6	641.241	94.681	695.275	-13653694.00	1452141.00	-12875597.00	-17804553.00	2281188.00	4445193.00	
7	641.203	96.218	695.357	-13653692.99	1452138.65	-12875593.35	-17804557.04	2281190.39	4445195.48	
7	641.242	94.681	695.275	-13653694.00	1452141.00	-12875597.00	-17804553.00	2281188.00	4445193.00	
8	641.546	96.843	695.346	-13653691.25	1452137.61	-12875590.03	-17804559.00	2281191.67	4445196.88	
8	641.243	94.680	695.274	-13653694.00	1452141.00	-12875597.00	-17804553.00	2281188.00	4445193.00	
9	641.597	96.951	695.294	-13653691.03	1452137.42	-12875589.43	-17804559.07	2281192.14	4445197.38	
9	641.243	94.680	695.274	-13653694.00	1452141.00	-12875597.00	-17804553.00	2281188.00	4445193.00	
10	641.413	96.573	695.260	-13653692.03	1452138.03	-12875591.27	-17804557.71	2281191.59	4445196.76	
10	641.243	94.680	695.274	-13653694.00	1452141.00	-12875597.00	-17804553.00	2281188.00	4445193.00	
11	641.452	96.602	695.280	-13653691.84	1452137.98	-12875591.03	-17804557.94	2281191.55	4445196.73	
11	640.371	96.696	695.178	-13653696.00	1452138.00	-12875596.00	-17804556.00	2281192.00	4445197.00	
12	641.422	96.541	695.293	-13653691.99	1452138.09	-12875591.35	-17804557.81	2281191.36	4445196.52	
12	641.244	94.678	695.274	-13653694.00	1452141.00	-12875597.00	-17804553.00	2281188.00	4445193.00	
13	641.580	96.830	695.308	-13653691.17	1452137.62	-12875589.84	-17804558.82	2281191.85	4445197.06	
13	641.244	94.678	695.273	-13653694.00	1452141.00	-12875597.00	-17804553.00	2281188.00	4445193.00	
14	641.574	96.828	695.324	-13653691.18	1452137.63	-12875589.90	-17804558.89	2281191.76	4445196.97	
14	641.245	94.677	695.273	-13653694.00	1452141.00	-12875597.00	-17804553.00	2281188.00	4445193.00	
15	641.516	96.757	695.328	-13653691.44	1452137.75	-12875590.37	-17804558.66	2281191.59	4445196.79	
15	640.998	95.965	695.328	-13653694.00	1452139.00	-12875595.00	-17804556.00	2281190.00	4445195.00	
16	641.559	96.921	695.349	-13653691.14	1452137.49	-12875589.76	-17804559.22	2281191.80	4445197.01	
16	640.998	95.964	695.328	-13653694.00	1452139.00	-12875595.00	-17804556.00	2281190.00	4445195.00	
.
.
.
150	640.991	95.973	695.332	-13653693.95	1452139.03	-12875594.96	-17804555.98	2281190.08	4445195.06	-17088833.01
150	640.995	95.965	695.333	-13653694.00	1452139.00	-12875595.00	-17804556.00	2281190.00	4445195.00	-17088833.00
151	640.991	95.972	695.332	-13653693.95	1452139.03	-12875594.96	-17804555.98	2281190.08	4445195.06	-17088833.01
151	640.995	95.965	695.333	-13653694.00	1452139.00	-12875595.00	-17804556.00	2281190.00	4445195.00	-17088833.00
152	640.991	95.972	695.333	-13653693.95	1452139.03	-12875594.96	-17804555.98	2281190.08	4445195.06	-17088833.01
152	640.995	95.965	695.333	-13653694.00	1452139.00	-12875595.00	-17804556.00	2281190.00	4445195.00	-17088833.00
153	640.991	95.972	695.333	-13653693.95	1452139.03	-12875594.96	-17804555.98	2281190.08	4445195.06	-17088833.01
153	640.995	95.965	695.333	-13653694.00	1452139.00	-12875595.00	-17804556.00	2281190.00	4445195.00	-17088833.00
154	640.992	95.973	695.334	-13653693.95	1452139.03	-12875594.96	-17804555.99	2281190.08	4445195.06	-17088833.01
154	640.995	95.965	695.333	-13653694.00	1452139.00	-12875595.00	-17804556.00	2281190.00	4445195.00	-17088833.00
155	640.992	95.974	695.335	-13653693.94	1452139.03	-12875594.95	-17804555.99	2281190.07	4445195.06	-17088833.01
155	640.995	95.965	695.333	-13653694.00	1452139.00	-12875595.00	-17804556.00	2281190.00	4445195.00	-17088833.00
156	640.993	95.975	695.335	-13653693.94	1452139.03	-12875594.95	-17804556.00	2281190.07	4445195.06	-17088833.01
156	640.995	95.965	695.333	-13653694.00	1452139.00	-12875595.00	-17804556.00	2281190.00	4445195.00	-17088833.00
157	640.992	95.974	695.335	-13653693.94	1452139.03	-12875594.95	-17804555.99	2281190.07	4445195.05	-17088833.02
157	640.995	95.964	695.333	-13653694.00	1452139.00	-12875595.00	-17804556.00	2281190.00	4445195.00	-17088833.00

- Obs. :
- 1) cada época apresenta em primeiro a solução flutuante (real) e, em segundo a solução com as ambigüidades inteiras;
 - 2) os valores inteiros das ambigüidades só se estabilizaram na 15ª época (3 minutos e 45 segundos)
 - 3) os valores obtidos pelo programa elaborado, para essa linha de base, resultaram em uma diferença de 3,1 mm em relação aos obtidos pelo *software* Prism da Ashtech, veja a página seguinte.

Linha de Base 3 (P003 – PARA) : comprimento 950,564 5 m

Estação fixa: **P003**

$\phi = 25^{\circ} 27' 19,35874''$ S

$\lambda = 49^{\circ} 14' 11,05371''$ W

H = 912,357 4 m

Extrato dos resultados do processamento no *software Prism*

THE FIXED DOUBLE DIFFERENCE SOLUTION (L1)
 Measure of geometry: 0.012487 Wavelength = 0.190294 (m/cycle)
 num_meas = 1055 num_used = 1028 rms_resid = 0.011816(m)
 Post-Fit Chisq = 4053.124 NDF = 7.139

Reference SV: 5 Integer Search Ratio = 100.000

SV	Ambiguity	FIT	Meas	SV	Ambiguity	FIT	Meas
6	12875595.000X	0.039	160	8	-2281190.000X	0.095	141
9	-4445195.000X	0.085	152	10	17804556.000X	0.043	160
17	17088833.000X	0.086	95	24	-1452139.000X	0.033	160
30	13653694.000X	0.034	160				

Sigmax (m): 0.013627
 Sigmay (m): 0.016253
 Sigmaz (m): 0.006939

x y z
 x 1.00
 y-0.83y 1.00
 z-0.64z 0.65z 1.00

del_station: -0.000184 0.000233 -0.000011

Station1: FIXED STATION			Station2: UNKNOWN STATION		
(00000)	(P003)		(00000)	(PARA)	
Latitude: -25.45537743	-25 27 19.35874		-25.44836914	-25 26 54.12892	
E-Long : 310.76359619	310 45 48.94629		310.76904532	310 46 8.56317	
W-Long : 49.23640381	49 14 11.05371		49.23095468	49 13 51.43683	
E-Height: 912.3574			925.8392		

Baseline vector: 640.9955 95.9668 695.3331

Mark1_xyz : 3763110.6927 -4365209.7915 -2725100.1167
 Az1 El1 D1 : 35.22031 0.8084 950.5645
 E1 N1 U1 : 548.1240 776.4873 13.4818
 Mark2_xyz : 3763751.6882 -4365113.8247 -2724404.7836
 Az2 El2 D2 : 215.21797 -0.8169 950.5645
 E2 N2 U2 : -548.1569 -776.4882 -13.4818

Tue May 29 16:21:15 2001

Obs.: Não usar as coordenadas acima para outros trabalhos, porque no processamento não estão sendo consideradas as alturas das antenas. O intuito é de apenas mostrar as duplas diferenças de ambigüidades, e as componentes da linha de base.
